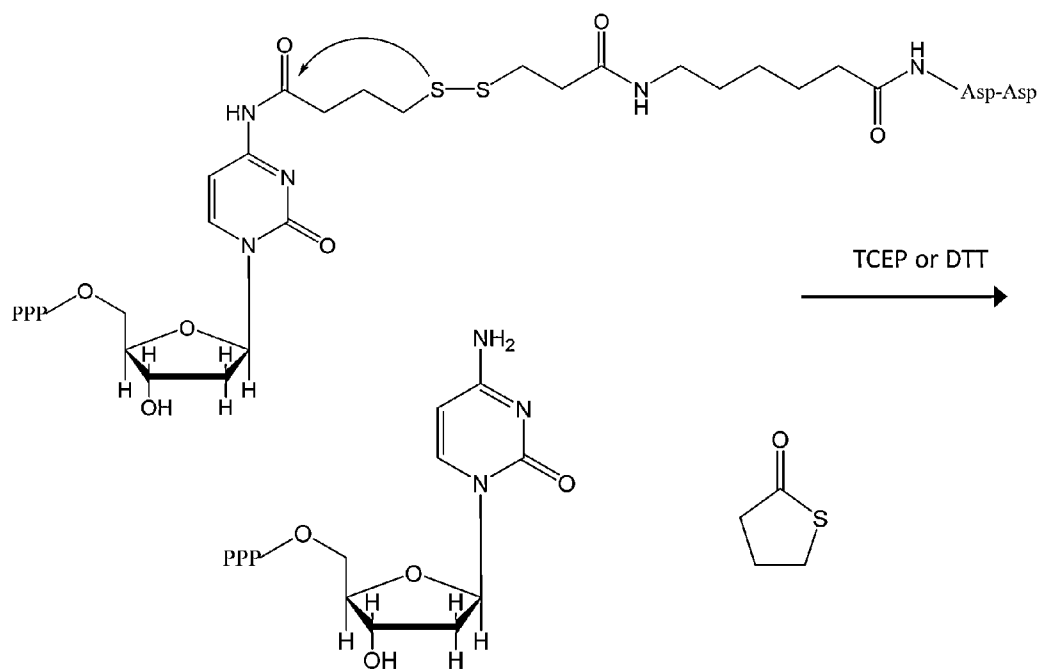
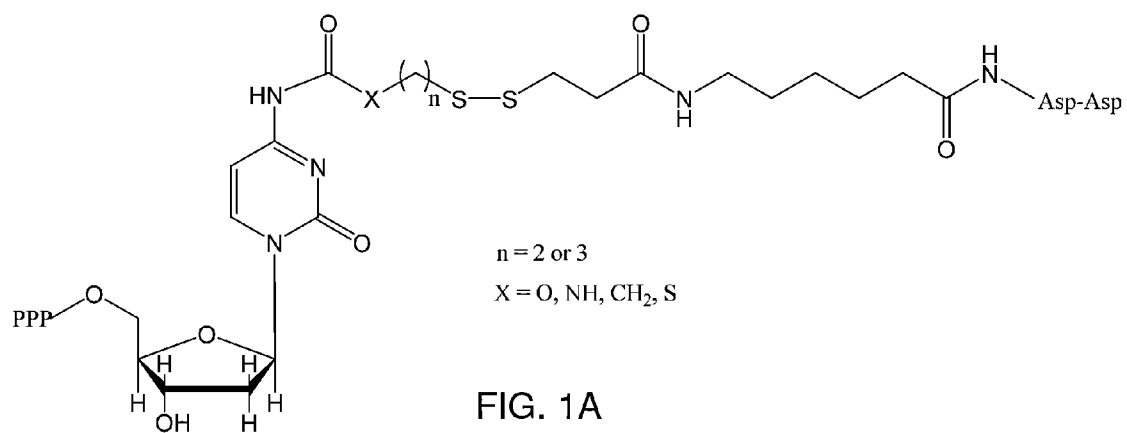




US 20140363852A1

(19) **United States**(12) **Patent Application Publication**
Efcavitch et al.(10) **Pub. No.: US 2014/0363852 A1**(43) **Pub. Date: Dec. 11, 2014**(54) **METHODS AND APPARATUS FOR
SYNTHESIZING NUCLEIC ACIDS****Publication Classification**(71) Applicant: **Molecular Assemblies, Inc.**, San Carlos,
CA (US)(51) **Int. Cl.**
C12P 19/34 (2006.01)(72) Inventors: **J. William Efcavitch**, San Carlos, CA
(US); **Suhaib Siddiqi**, Burlington, MA
(US)(52) **U.S. Cl.**
CPC **C12P 19/34** (2013.01)
USPC **435/91.5**(21) Appl. No.: **14/459,014**(57) **ABSTRACT**(22) Filed: **Aug. 13, 2014****Related U.S. Application Data**(63) Continuation-in-part of application No. 14/056,687,
filed on Oct. 17, 2013, now Pat. No. 8,808,989.(60) Provisional application No. 61/891,162, filed on Oct.
15, 2013, provisional application No. 61/807,327,
filed on Apr. 2, 2013.

The invention provides improved methods for synthesizing polynucleotides, such as DNA and RNA, using enzymes and specially designed nucleotide analogs. Using the methods of the invention, specific sequences of polynucleotides can be synthesized de novo, base by base, in an aqueous environment, without the use of a nucleic acid template. Because the nucleotide analogs have an unmodified 3' OH, i.e., as found in "natural" deoxyribose and ribose molecules, the analogs result in natural polynucleotides suitable for incorporation into biological systems.



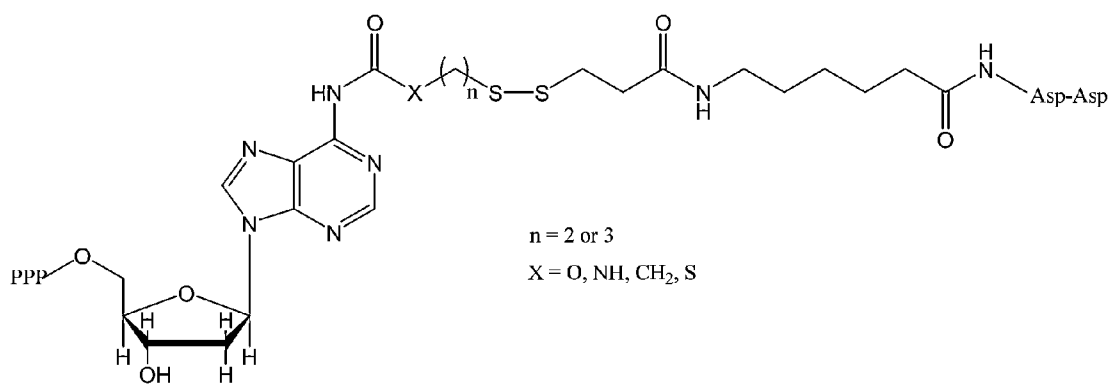


FIG. 2A

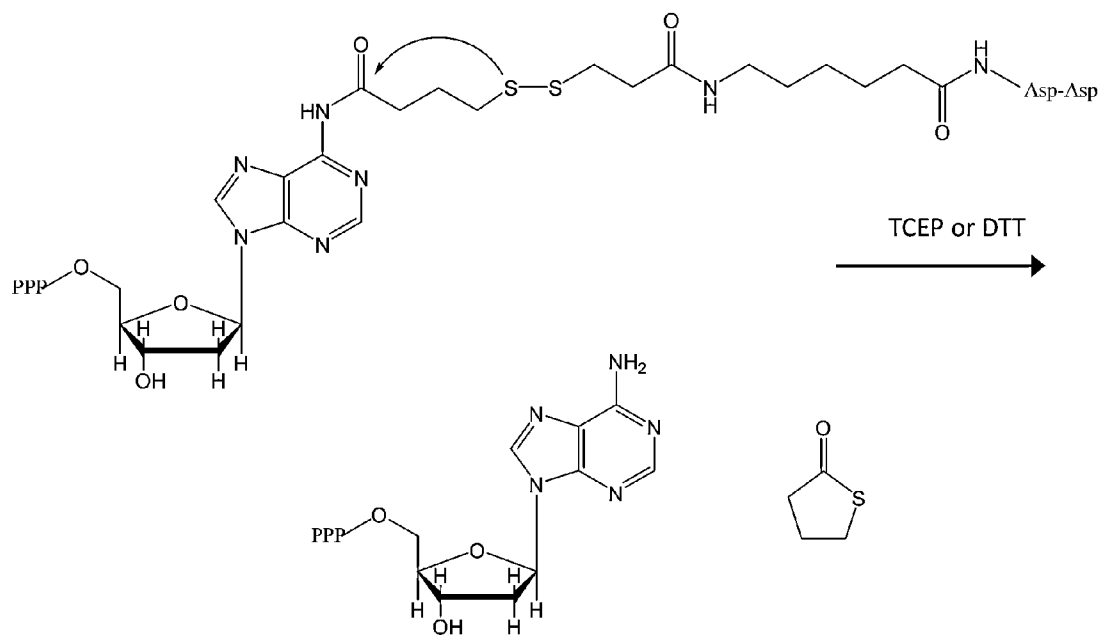


FIG. 2B

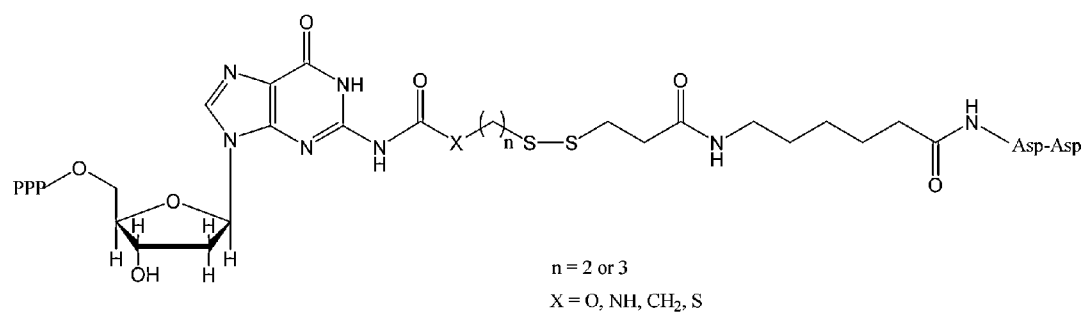


FIG. 3A

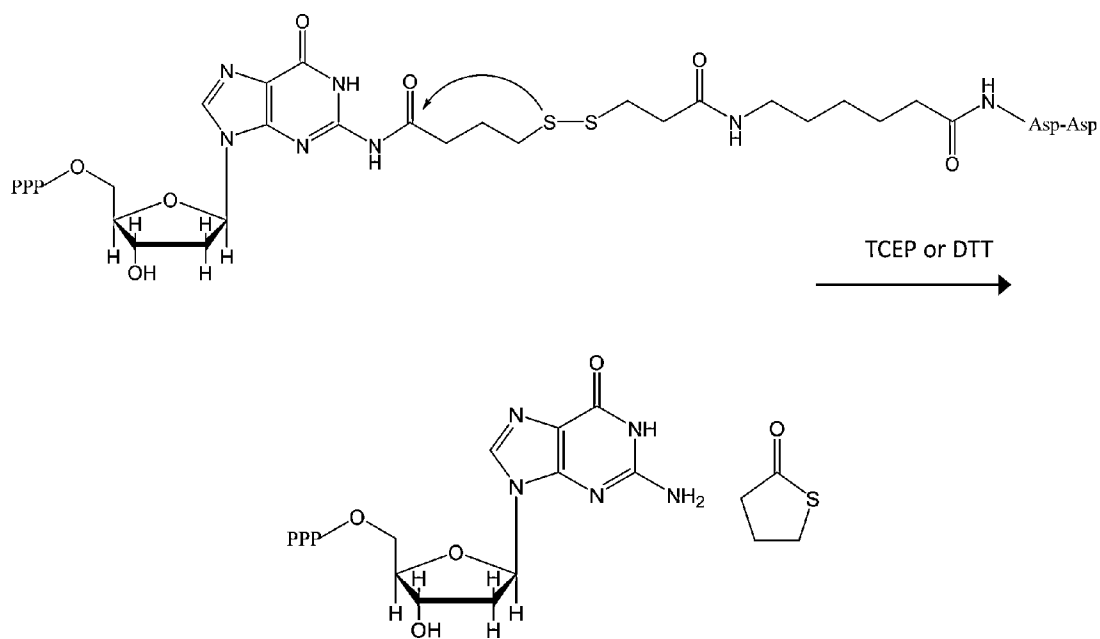


FIG. 3B

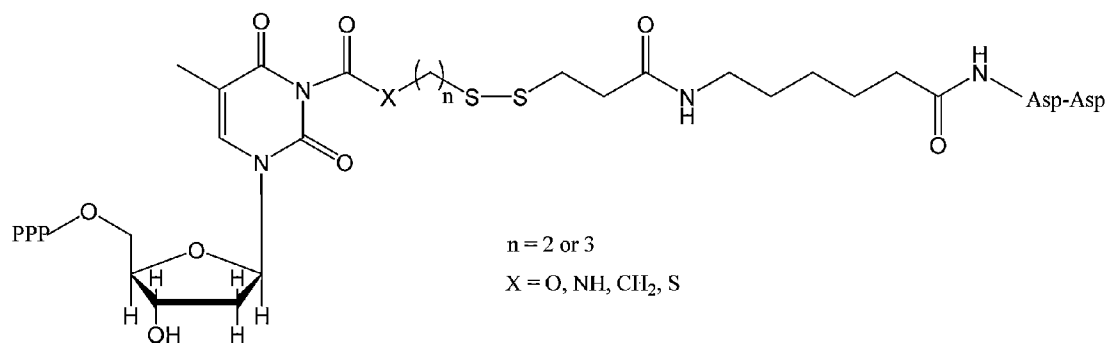


FIG. 4A

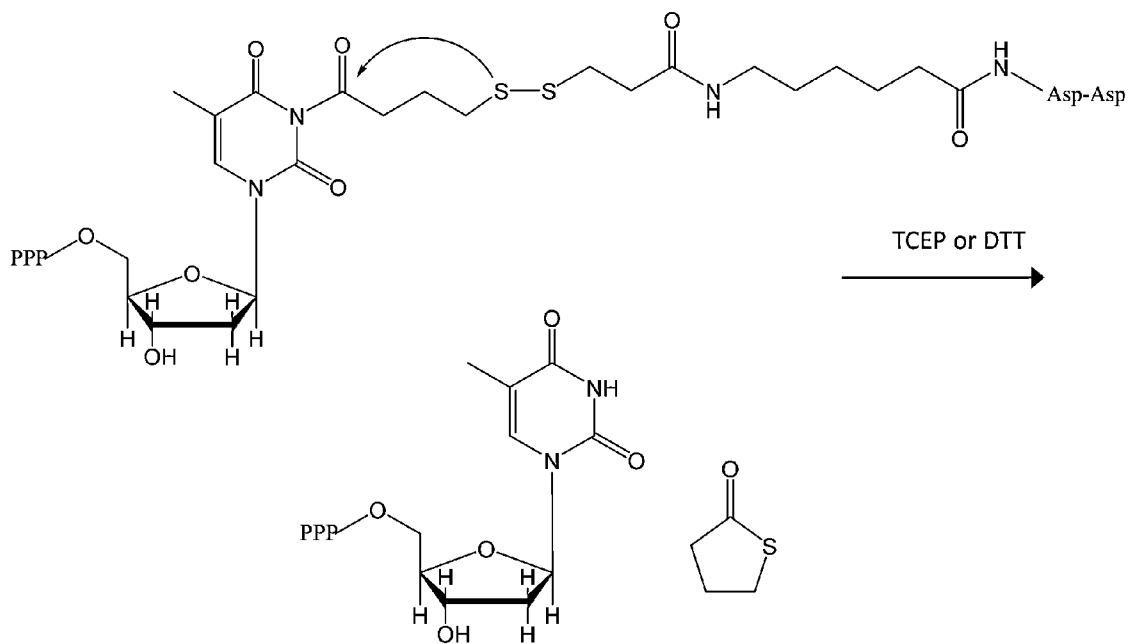


FIG. 4B

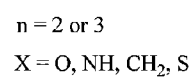


FIG. 5A

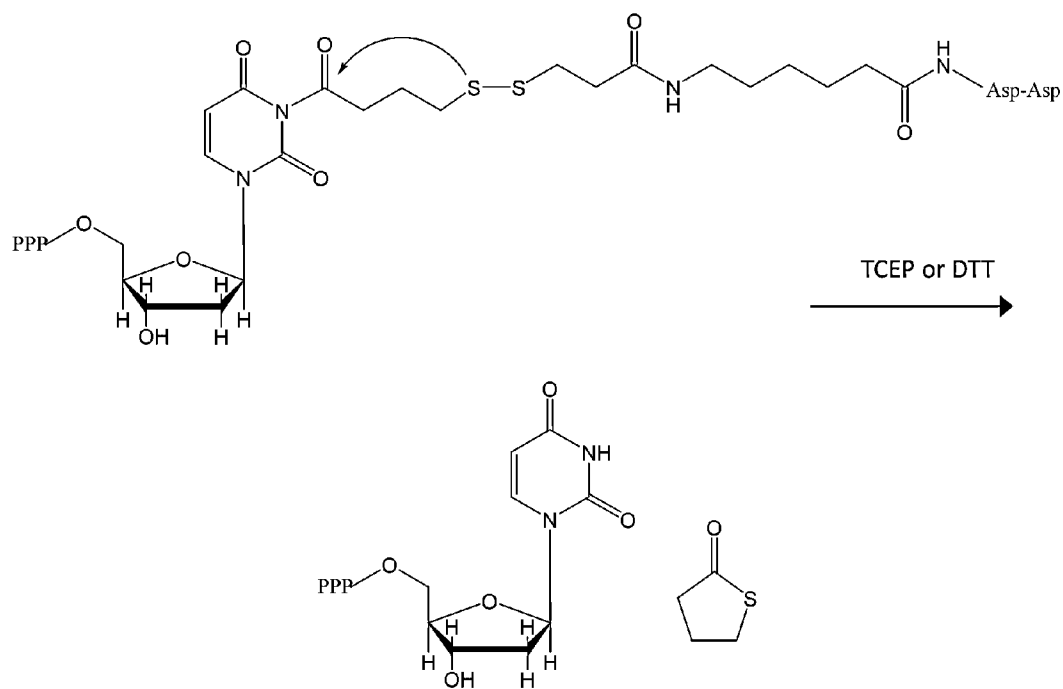


FIG. 5B

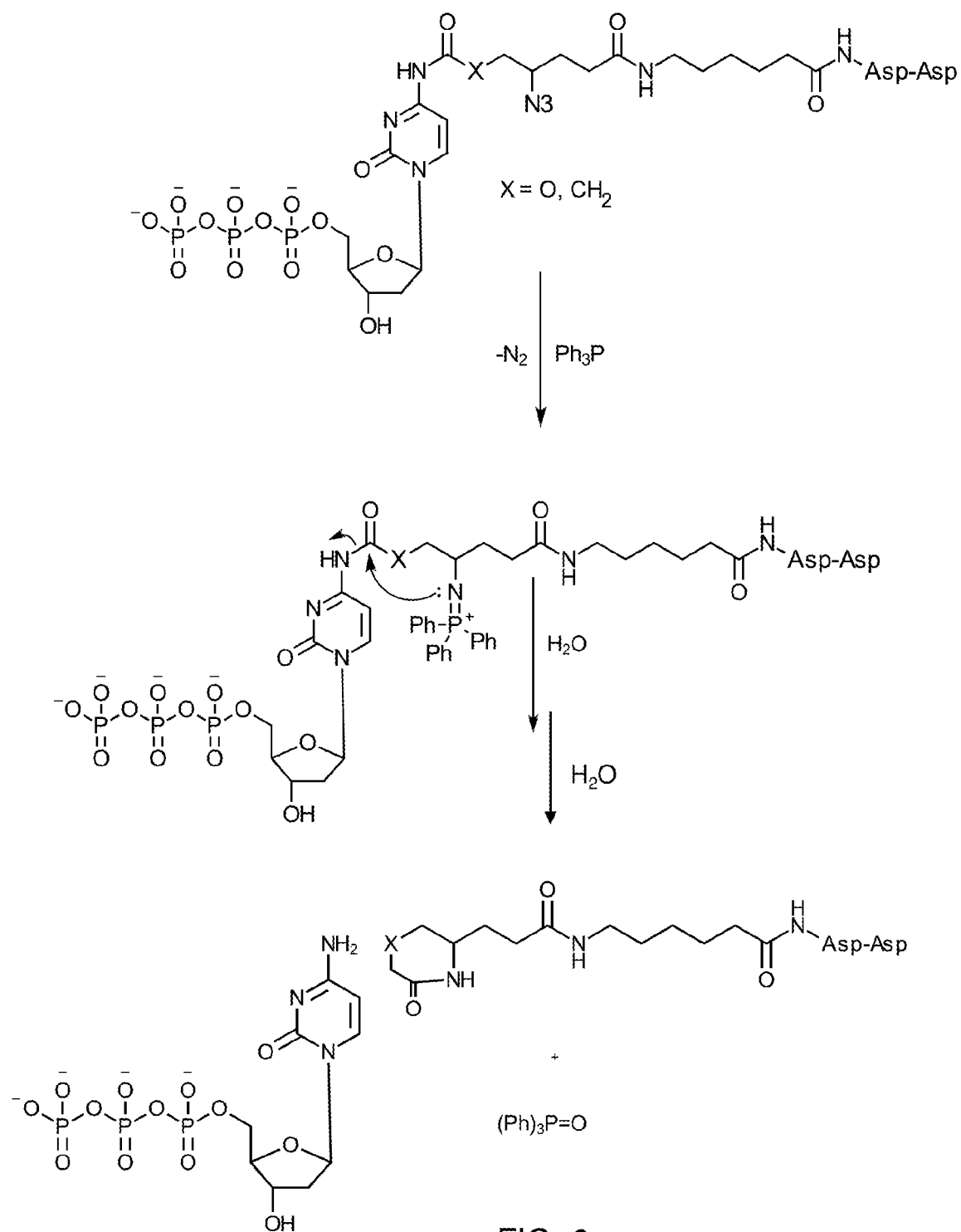
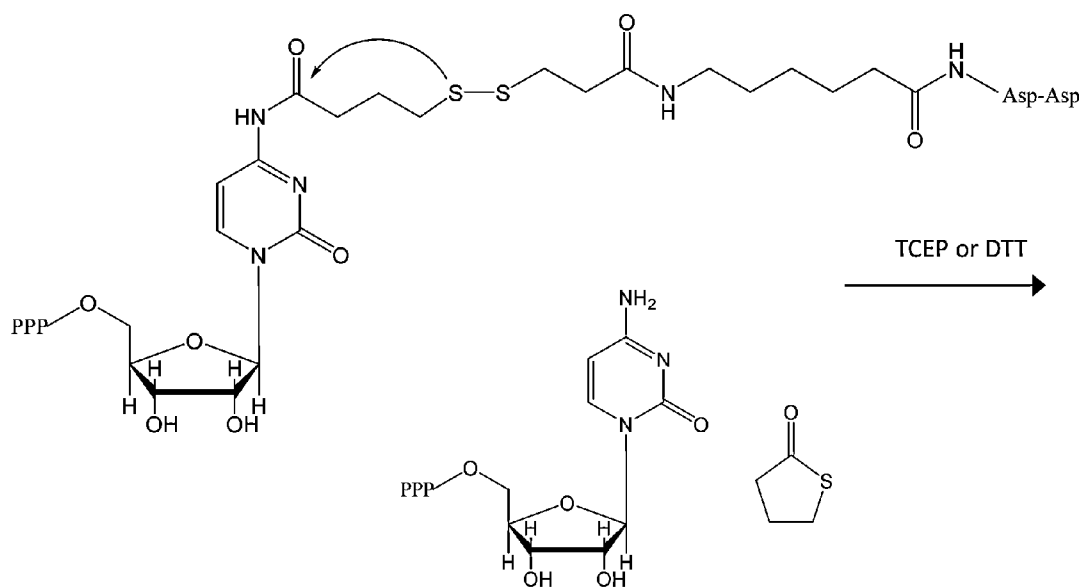
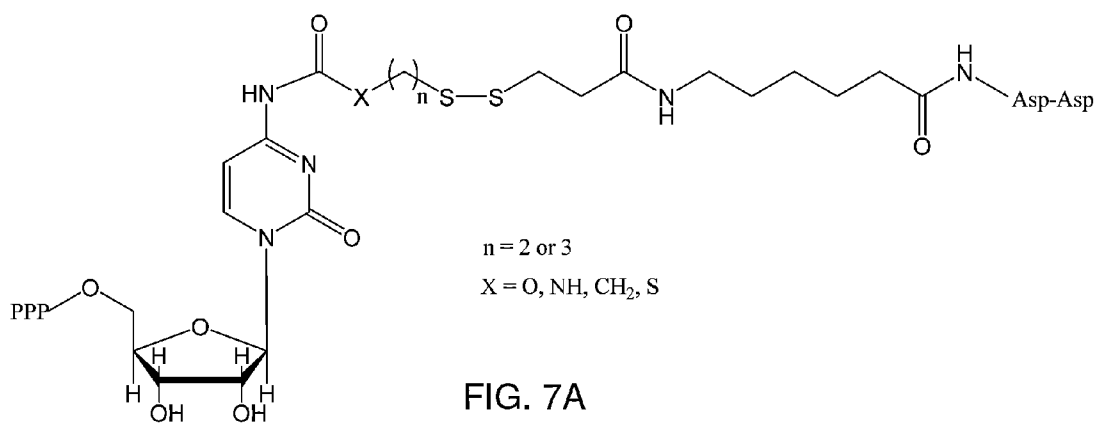


FIG. 6



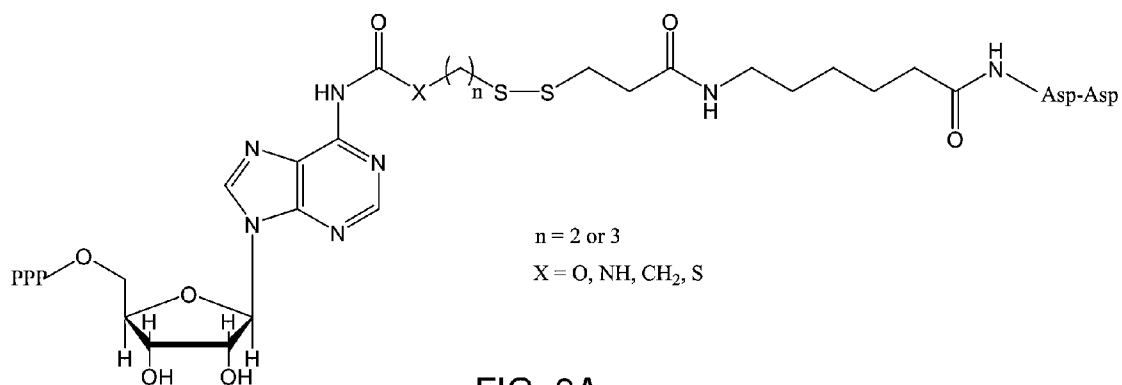


FIG. 8A

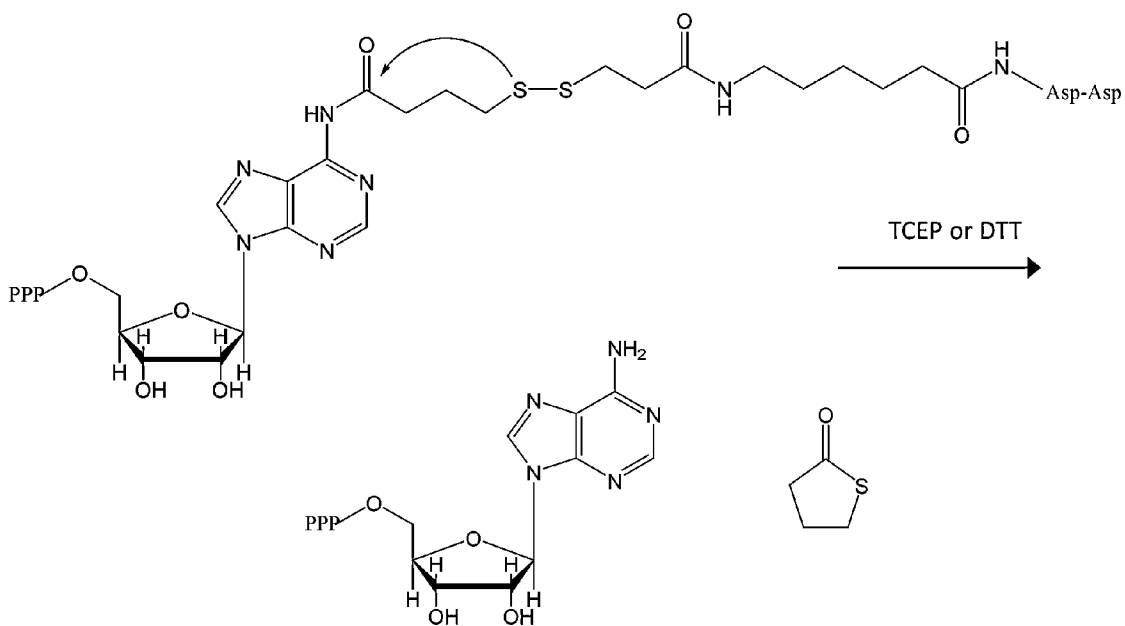


FIG. 8B

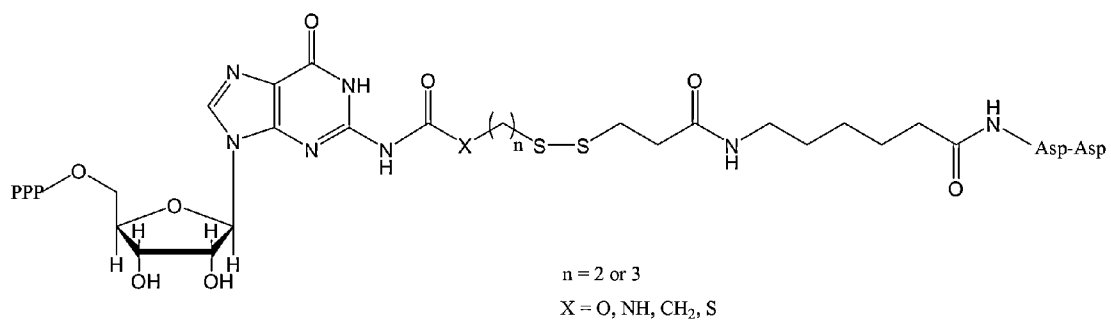


FIG. 9A

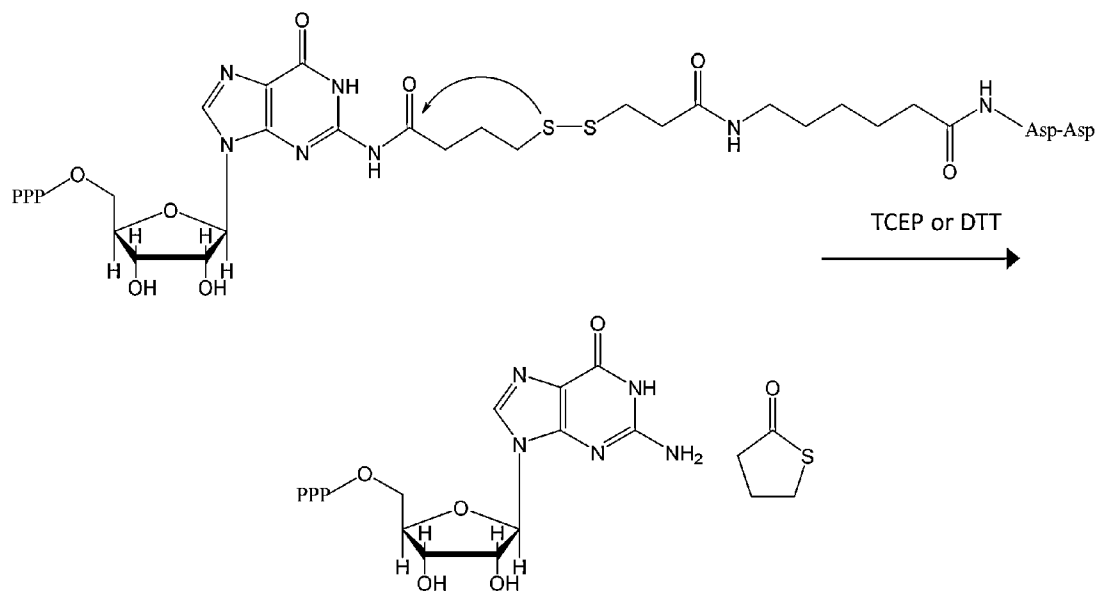
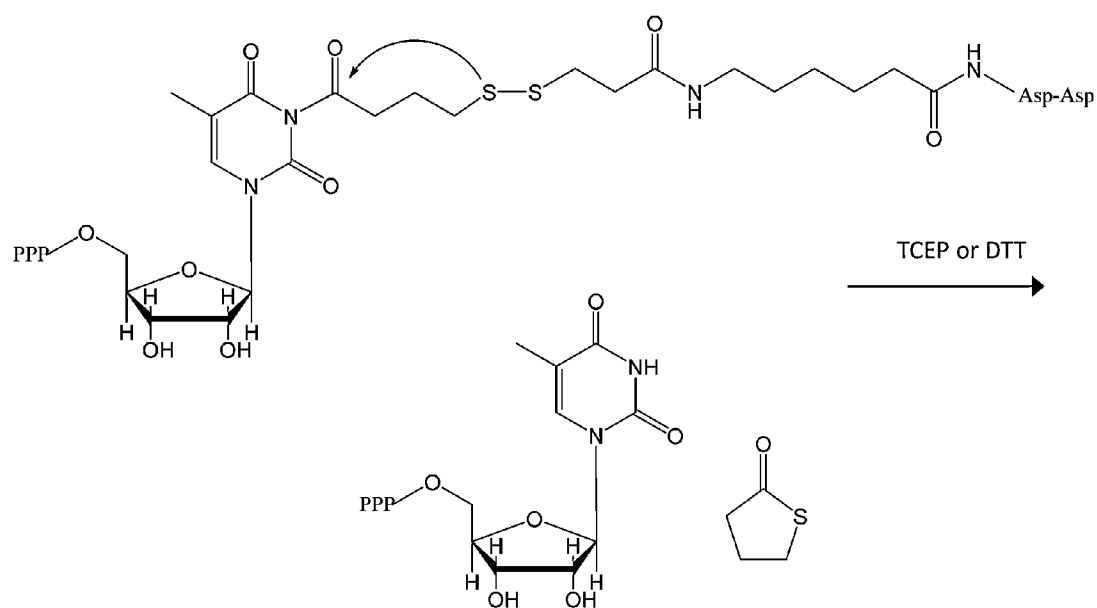
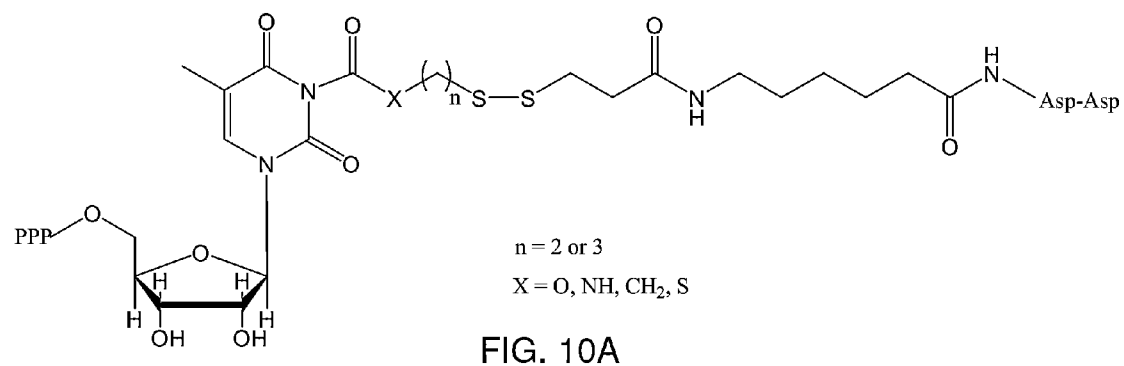


FIG. 9B



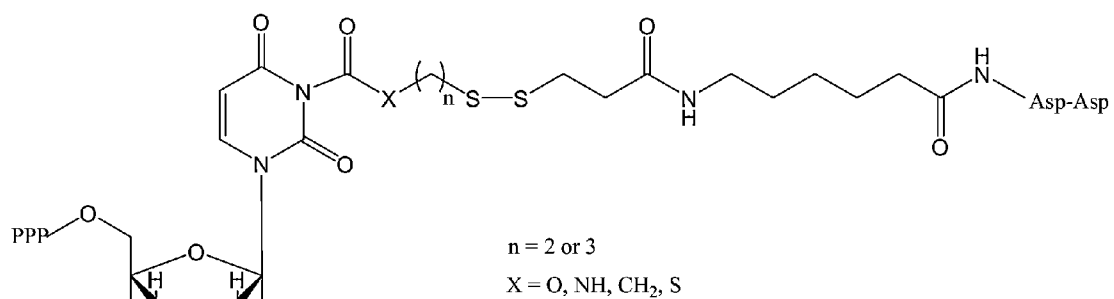


FIG. 11A

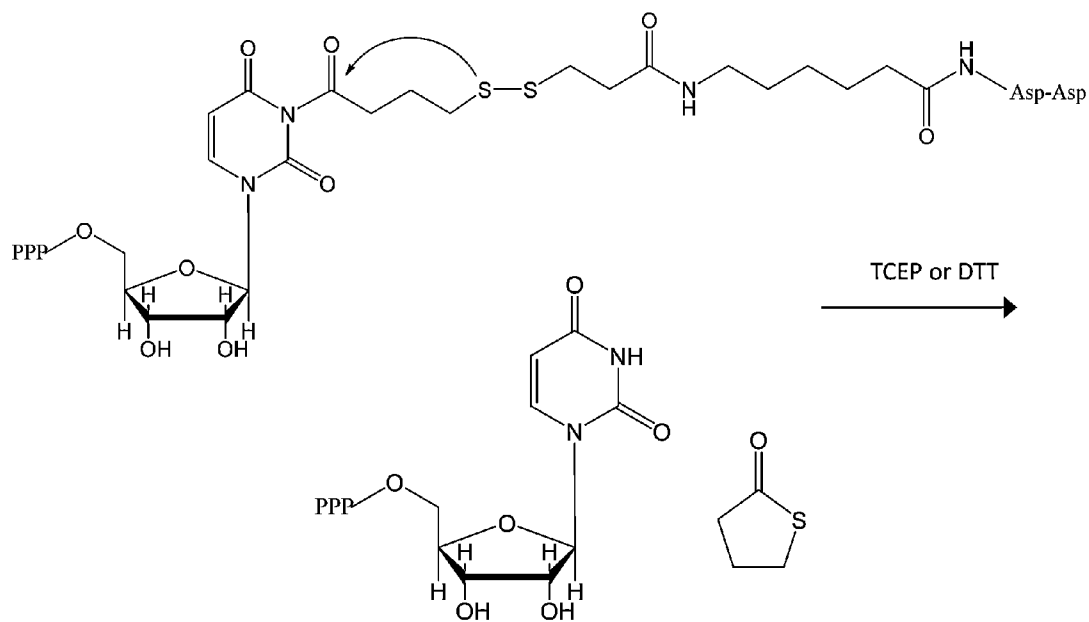


FIG. 11B

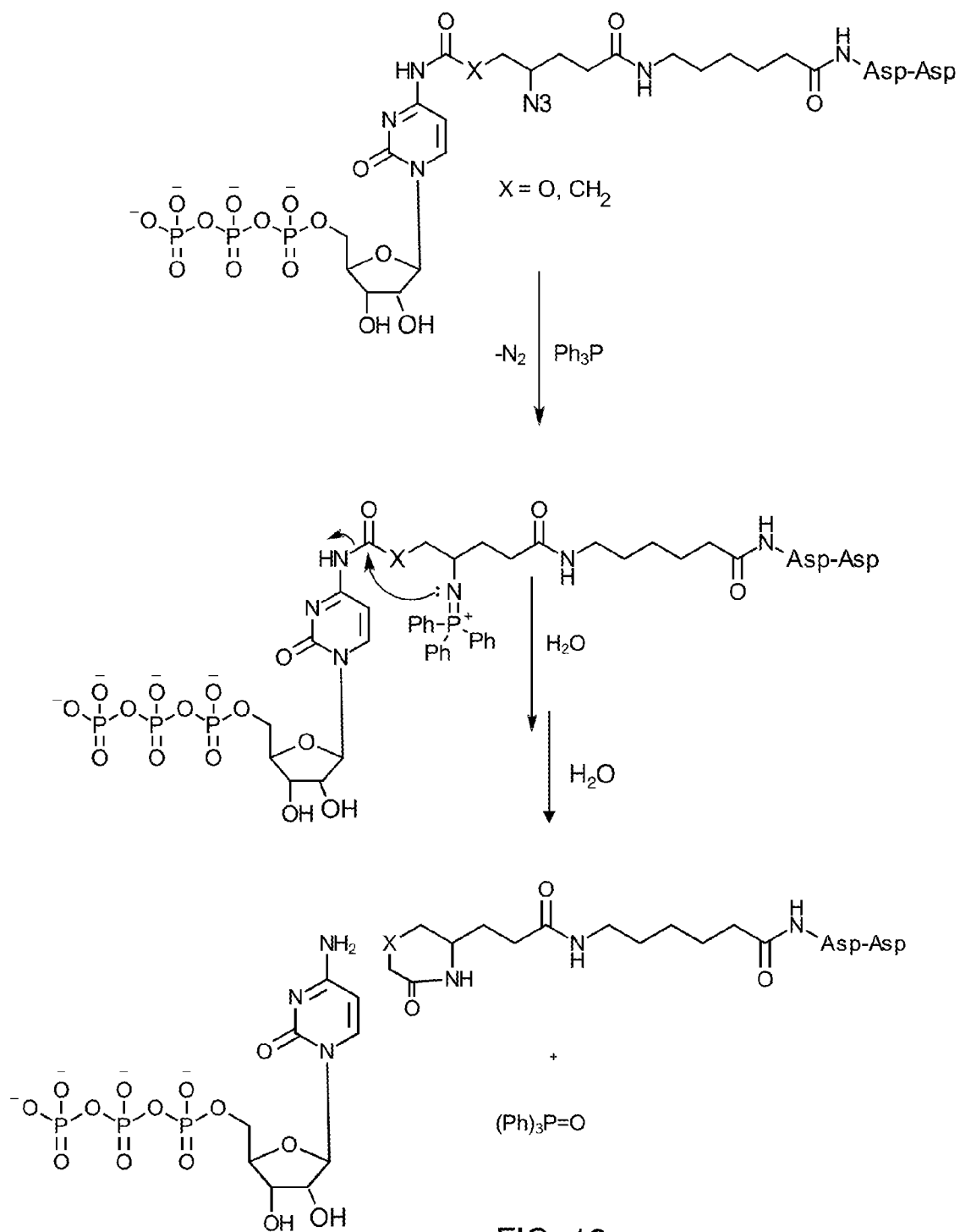


FIG. 12

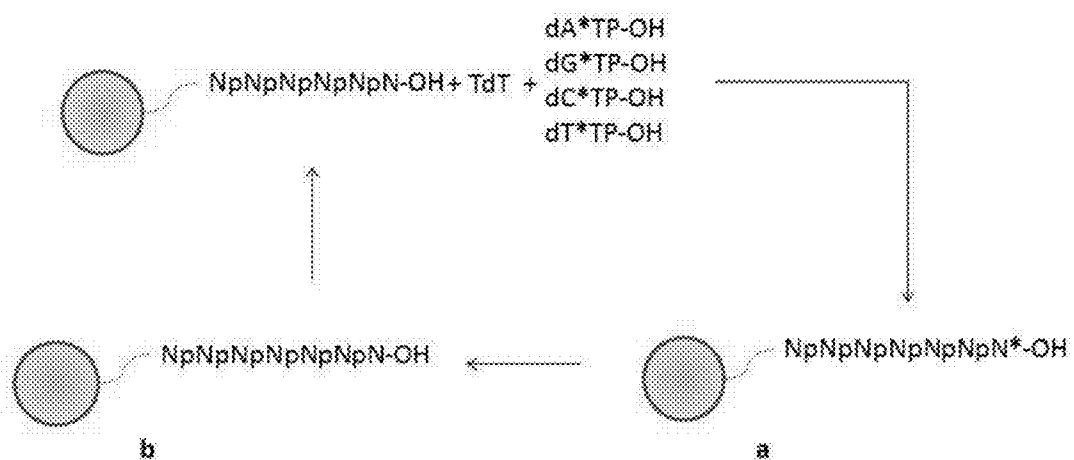
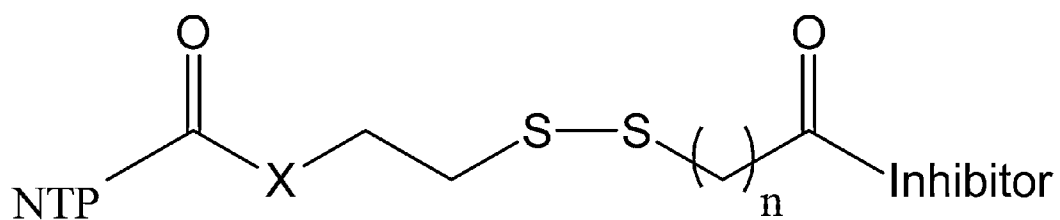


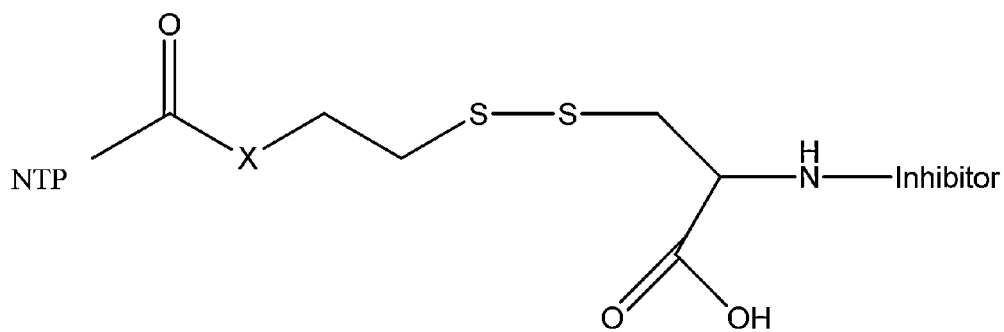
FIG. 13



$n = 1 - 4$

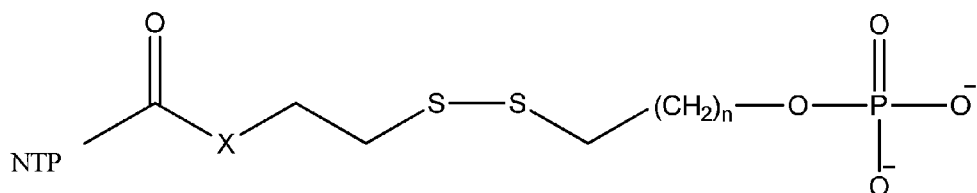
$\text{X} = \text{O}, \text{S}, \text{NH}, \text{CH}_2$

FIG. 14



$\text{X} = \text{O}, \text{S}, \text{NH}, \text{CH}_2$

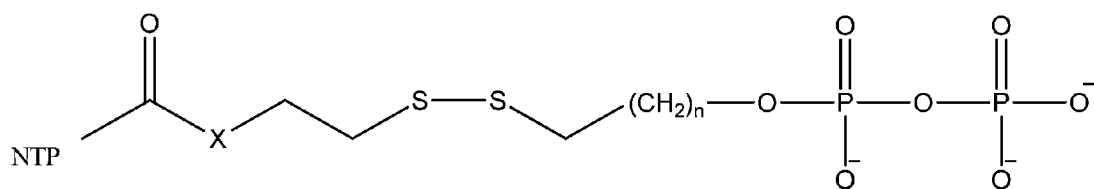
FIG. 15



$n = 1 - 4$

$\text{X} = \text{O}, \text{S}, \text{NH}, \text{CH}_2$

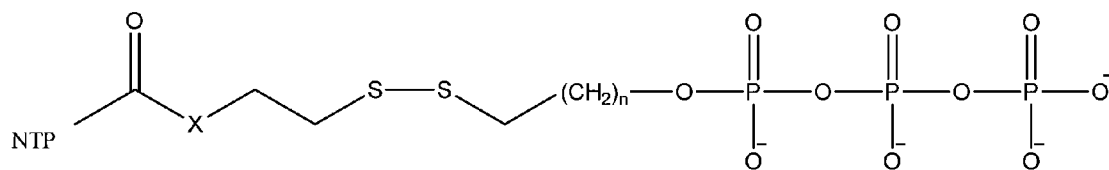
FIG. 16A



$n = 1 - 4$

$\text{X} = \text{O}, \text{S}, \text{NH}, \text{CH}_2$

FIG. 16B



$n = 1 - 4$

$\text{X} = \text{O}, \text{S}, \text{NH}, \text{CH}_2$

FIG. 16C

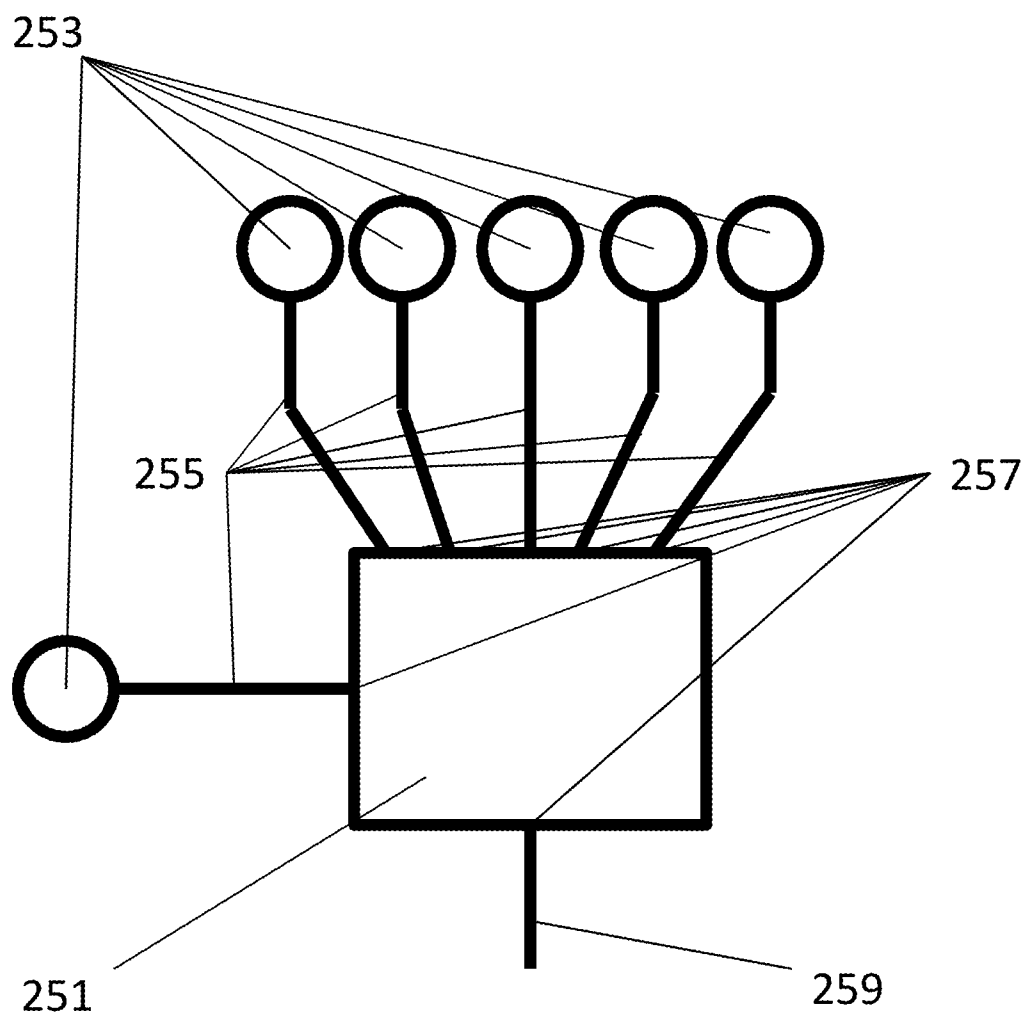


FIG. 17

METHODS AND APPARATUS FOR SYNTHESIZING NUCLEIC ACIDS

RELATED APPLICATIONS

[0001] This application is a continuation-in-part of U.S. application Ser. No. 14/056,687, filed on Oct. 17, 2013 which claims priority to U.S. Provisional Application Nos. 61/807,327, filed Apr. 2, 2013, and 61/891,162, filed Oct. 15, 2013, all of which are incorporated by reference in their entireties.

FIELD OF THE INVENTION

[0002] The invention relates to methods and apparatus for synthesizing polynucleotides (de novo) with a desired sequence and without the need for a template. As such, the invention provides the capacity to make libraries of polynucleotides of varying sequence and varying length for research, genetic engineering, and gene therapy.

BACKGROUND

[0003] Genetic engineering requires tools for determining the content of genetic material as well as tools for constructing desired genetic materials. The tools for determining the content of genetic material have made it possible to sequence an entire human genome in about one day for under \$1,000. (See Life Technologies, *Press Release: Benchtop Ion Proton™ Sequencer*, Jan. 10, 2012). In contrast, the tools for constructing desired genetic materials, e.g., de novo DNA synthesis, have not improved at the same pace. As a point of reference, over the past 25 years, the cost (per base) of de novo small nucleic acid synthesis has dropped 10-fold, while the cost (per base) of nucleic acid sequencing has dropped over 10,000,000-fold. The lack of progress in DNA synthesis now limits the pace of translational genomics, i.e., whereby the role of individual sequence variations are determined and used to develop therapeutic treatments.

[0004] Currently, most de novo nucleic acid sequences are synthesized using solid phase phosphoramidite-techniques developed more than 30 years ago. The technique involves the sequential de-protection and synthesis of sequences built from phosphoramidite reagents corresponding to natural (or non-natural) nucleic acid bases. Phosphoramidite nucleic acid synthesis is length-limited, however, in that nucleic acids greater than 200 base pairs (bp) in length experience high rates of breakage and side reactions. Additionally, phosphoramidite synthesis produces toxic by-products, and the disposal of this waste limits the availability of nucleic acid synthesizers, and increases the costs of contract oligo production. (It is estimated that the annual demand for oligonucleotide synthesis is responsible for greater than 300,000 gallons of hazardous chemical waste, including acetonitrile, trichloroacetic acid, toluene, tetrahydrofuran, and pyridine. See LeProust et al., *Nucleic Acids Res.*, vol. 38(8), p. 2522-2540, (2010), incorporated by reference herein in its entirety). Thus, there is a need for more efficient and cost-effective methods for oligonucleotide synthesis.

SUMMARY

[0005] The invention provides improved methods for nucleic acid synthesis. Methods of the invention provide faster and longer de novo synthesis of polynucleotides. As such, the invention dramatically reduces the overall cost of synthesizing custom nucleic acids. Methods of the invention are directed to template-independent synthesis of polynucle-

otides by using a nucleotidyl transferase enzyme to incorporate nucleotide analogs having an unmodified 3' hydroxyl coupled to an inhibitor by a cleavable linker. Because of the inhibitor, synthesis pauses with the addition of each new base, whereupon the linker is cleaved, separating the inhibitor and leaving a polynucleotide that is essentially identical to a naturally occurring nucleotide (i.e., is recognized by the enzyme as a substrate for further nucleotide incorporation).

[0006] The invention additionally includes an apparatus that utilizes methods of the invention for the production of custom polynucleotides. An apparatus of the invention includes one or more bioreactors providing aqueous conditions and a plurality of sources of nucleotide analogs. The bioreactor may be e.g., a reservoir, a flow cell, or a multi-well plate. Starting from a solid support, the polynucleotides are grown in the reactor by adding successive nucleotides via the natural activity of a nucleotidyl transferase, e.g., a terminal deoxynucleotidyl transferase (TdT) or any other enzyme which elongates DNA or RNA strands without template direction. Upon cleavage of the linker, a natural polynucleotide is exposed on the solid support. Once the sequence is complete, the support is cleaved away, leaving a polynucleotide essentially equivalent to that found in nature. In some embodiments, the apparatus is designed to recycle nucleotide analog solutions by recovering the solutions after nucleotide addition and reusing solutions for subsequent nucleotide addition. Thus, less waste is produced, and the overall cost per base is reduced as compared to state-of-the-art methods. In certain embodiments, a bioreactor may include a microfluidic device and/or use inkjet printing technology.

[0007] Other aspects of the invention are apparent to the skilled artisan upon consideration of the following figures and detailed description.

BRIEF DESCRIPTION OF THE DRAWINGS

[0008] FIG. 1A shows a genus of deoxycytidine triphosphate (dCTP) analogs having a cleavable terminator linked at the N-4 position;

[0009] FIG. 1B shows cleavage of the cleavable terminator from a dCTP analog of FIG. 1A to achieve a "natural" dCTP and a cyclic leaving molecule;

[0010] FIG. 2A shows a genus of deoxyadenosine triphosphate (dATP) analogs having a cleavable terminator linked at the N-6 position;

[0011] FIG. 2B shows cleavage of the cleavable terminator from a dATP analog of FIG. 2A to achieve a "natural" dATP and a cyclic leaving molecule;

[0012] FIG. 3A shows a genus of deoxyguanosine triphosphate (dGTP) analogs having a cleavable terminator linked at the N-2 position;

[0013] FIG. 3B shows cleavage of the cleavable terminator from a dGTP analog of FIG. 3A to achieve a "natural" dGTP and a cyclic leaving molecule;

[0014] FIG. 4A shows a genus of deoxythymidine triphosphate (dTTP) analogs having a cleavable terminator linked at the N-3 position;

[0015] FIG. 4B shows cleavage of the cleavable terminator from a dTTP analog of FIG. 4A to achieve a "natural" dTTP and a cyclic leaving molecule;

[0016] FIG. 5A shows a genus of deoxyuridine triphosphate (dUTP) analogs having a cleavable terminator linked at the N-3 position;

[0017] FIG. 5B shows cleavage of the cleavable terminator from a dUTP analog of FIG. 5A to achieve a dUTP and a cyclic leaving molecule;

[0018] FIG. 6 shows an exemplary deoxycytidine triphosphate (dCTP) analog having a Staudinger linker connecting a blocking Asp-Asp molecule to the N-4 position of the deoxycytidine and subsequent cleavage of the Staudinger linker under aqueous conditions to achieve a dCTP and a leaving group;

[0019] FIG. 7A shows a genus of cytidine triphosphate (rCTP) analogs having a cleavable terminator linked at the N-4 position;

[0020] FIG. 7B shows cleavage of the cleavable terminator from a rCTP analog of FIG. 7A to achieve a "natural" rCTP and a cyclic leaving molecule;

[0021] FIG. 8A shows a genus of adenosine triphosphate (rATP) analogs having a cleavable terminator linked at the N-6 position;

[0022] FIG. 8B shows cleavage of the cleavable terminator from an rATP analog of FIG. 8A to achieve a "natural" rATP and a cyclic leaving molecule;

[0023] FIG. 9A shows a genus of guanosine triphosphate (rGTP) analogs having a cleavable terminator linked at the N-2 position;

[0024] FIG. 9B shows cleavage of the cleavable terminator from a rGTP analog of FIG. 9A to achieve a "natural" rGTP and a cyclic leaving molecule;

[0025] FIG. 10A shows a genus of thymidine triphosphate (rTTP) analogs having a cleavable terminator linked at the N-3 position;

[0026] FIG. 10B shows cleavage of the cleavable terminator from a rTTP analog of FIG. 10A to achieve a "natural" rTTP and a cyclic leaving molecule;

[0027] FIG. 11A shows a genus of uridine triphosphate (rUTP) analogs having a cleavable terminator linked at the N-3 position;

[0028] FIG. 11B shows cleavage of the cleavable terminator from a rUTP analog of FIG. 11A to achieve a rUTP and a cyclic leaving molecule;

[0029] FIG. 12 shows an exemplary cytidine triphosphate (rCTP) analog having a Staudinger linker connecting a blocking Asp-Asp molecule to the N-4 position of the cytidine and subsequent cleavage of the Staudinger linker under aqueous conditions to achieve a rCTP and a leaving group;

[0030] FIG. 13 shows an exemplary terminal deoxynucleotidyl transferase (TdT) mediated polynucleotide synthetic cycle, including: (a) incorporation of a nucleotide triphosphate analog comprising cleavable terminator, dN*TP-OH, and (b) removal of the terminating blocking group (indicated by *), thus enabling the next dN*TP-OH to be incorporated, wherein N=A, G, C, or T;

[0031] FIG. 14 shows an exemplary nucleotide analog with a cleavable linker comprising a variable number of methylene bridges;

[0032] FIG. 15 shows an exemplary nucleotide analog with a cleavable linker comprising a cysteine residue;

[0033] FIG. 16A shows an exemplary nucleotide analog with an anionic inhibitor comprising a single phosphate group;

[0034] FIG. 16B shows an exemplary nucleotide analog with an anionic inhibitor comprising two phosphate groups;

[0035] FIG. 16C shows an exemplary nucleotide analog with an anionic inhibitor comprising three phosphate groups;

[0036] FIG. 17 shows an exemplary microfluidic polynucleotide synthesis device.

DETAILED DESCRIPTION

[0037] The invention provides improved methods for synthesizing polynucleotides, such as DNA and RNA, using enzymes and nucleic acid analogs. Using the disclosed methods, specific sequences of polynucleotides can be synthesized de novo, base by base, in an aqueous environment, without the use of a nucleic acid template. Additionally, because the nucleotide analogs have an unmodified 3' hydroxyls, i.e., as found in "natural" deoxyribose and ribose molecules, the analogs result in "natural" nucleotides when an inhibitor, attached via a cleavable linker, is removed from the base. Other nucleotide analogs can also be used which, for example, include self-eliminating linkers, or nucleotides with modified phosphate groups. In most instances, the blocking group is designed to not leave behind substantial additional molecules, i.e., designed to leave behind "scarless" nucleotides that are recognized as "natural" nucleotides by the enzyme. Thus, at the conclusion of the synthesis, upon removal of the last blocking group, the synthesized polynucleotide is chemically and structurally equivalent to the naturally-occurring polynucleotide with the same sequence. The synthetic polynucleotide can, thus, be incorporated into living systems without concern that the synthesized polynucleotide will interfere with biochemical pathways or metabolism.

[0038] The process and analogs of the current invention can be used for the non-templated enzymatic synthesis of useful oligo- and oligodeoxynucleotides especially of long oligonucleotides (<5000 nt). Products can be single strand or partially double strand depending upon the initiator used. The synthesis of long oligonucleotides requires high efficiency incorporation and high efficiency of reversible terminator removal. The initiator bound to the solid support consists of a short, single strand DNA sequence that is either a short piece of the user defined sequence or a universal initiator from which the user defined single strand product is removed.

[0039] In one aspect, the disclosed methods employ commercially-available nucleotidyl transferase enzymes, such as terminal deoxynucleotidyl transferase (TdT), to synthesize polynucleotides from nucleotide analogs in a step-by-step fashion. The nucleotide analogs are of the form:

[0040] NTP-linker-inhibitor

wherein NTP is a nucleotide triphosphate (i.e., a dNTP or an rNTP), the linker is a cleavable linker between the pyridine or pyrimidine of the base, and the inhibitor is a group that prevents the enzyme from incorporating subsequent nucleotides. At each step, a new nucleotide analog is incorporated into the growing polynucleotide chain, whereupon the enzyme is blocked from adding an additional nucleotide by the inhibitor group. Once the enzyme has stopped, the excess nucleotide analogs can be removed from the growing chain, the inhibitor can be cleaved from the NTP, and new nucleotide analogs can be introduced in order to add the next nucleotide to the chain. By repeating the steps sequentially, it is possible to quickly construct nucleotide sequences of a desired length and sequence. Advantages of using nucleotidyl transferases for polynucleotide synthesis include: 1) 3'-extension activity using single strand (ss) initiating primers in a template-independent polymerization, 2) the ability to extend primers in a highly efficient manner resulting in the addition of thousands of nucleotides, and 3) the acceptance of a wide variety of

modified and substituted NTPs as efficient substrates. In addition, the invention can make use of an initiator sequence that is a substrate for nucleotidyl transferase. The initiator is attached to a solid support and serves as a binding site for the enzyme. The initiator is preferably a universal initiator for the enzyme, such as a homopolymer sequence and is recyclable on the solid support, the formed oligonucleotide being cleavable from the initiator.

[0041] Methods of the invention are well-suited to a variety of applications that currently use synthetic nucleic acids, e.g., phosphoramidite-synthesized DNA oligos. For example, polynucleotides synthesized with the methods of the invention can be used as primers for nucleic acid amplification, hybridization probes for detection of specific markers, and for incorporation into plasmids for genetic engineering. However, because the disclosed methods produce longer synthetic strings of nucleotides, at a faster rate, and in an aqueous environment, the disclosed methods also lend themselves to high-throughput applications, such as screening for expression of genetic variation in cellular assays, as well as synthetic biology. Furthermore, the methods of the invention will provide the functionality needed for next-generation applications, such as using DNA as synthetic read/write memory, or creating macroscopic materials synthesized completely (or partially) from DNA.

[0042] The invention and systems described herein provide for synthesis of polynucleotides, including deoxyribonucleic acids (DNA) and ribonucleic acids (RNA). While synthetic pathways for “natural” nucleotides, such as DNA and RNA, are described in the context of the common nucleic acid bases, e.g., adenine (A), guanine (G), cytosine (C), thymine (T), and uracil (U), it is to be understood that the methods of the invention can be applied to so-called “non-natural” nucleotides, including nucleotides incorporating universal bases such as 3-nitropyrrole 2'-deoxynucleoside and 5-nitroindole 2'-deoxynucleoside, alpha phosphorothiolate, phosphorothioate nucleotide triphosphates, or purine or pyrimidine conjugates that have other desirable properties, such as fluorescence. Other examples of purine and pyrimidine bases include pyrazolo[3,4-d]pyrimidines, 5-methylcytosine (5-me-C), 5-hydroxymethyl cytosine, xanthine, hypoxanthine, 2-aminoadenine, 6-methyl and other alkyl derivatives of adenine and guanine, 2-propyl and other alkyl derivatives of adenine and guanine, 2-thiouracil, 2-thiothymine and 2-thiocytosine, 5-propynyl uracil and cytosine, 6-azo uracil, cytosine and thymine, 5-uracil (pseudouracil), 4-thiouracil, 8-halo (e.g., 8-bromo), 8-amino, 8-thiol, 8-thioalkyl, 8-hydroxyl and other 8-substituted adenines and guanines, 5-halo particularly 5-bromo, 5-trifluoromethyl and other 5-substituted uracils and cytosines, 7-methylguanine and 7-methyladenine, 8-azaguanine and 8-azaadenine, deazaguanine, 7-deazaguanine, 3-deazaguanine, deazaadenine, 7-deazaadenine, 3-deazaadenine, pyrazolo[3,4-d]pyrimidine, imidazo[1,5-a]1,3,5 triazinones, 9-deazapurines, imidazo[4,5-d]pyrazines, thiazolo [4,5-d]pyrimidines, pyrazin-2-ones, 1,2,4-triazine, pyridazine; and 1,3,5 triazine. In some instances, it may be useful to produce nucleotide sequences having unreactive, but approximately equivalent bases, i.e., bases that do not

react with other proteins, i.e., transcriptases, thus allowing the influence of sequence information to be decoupled from the structural effects of the bases.

Analogous

[0043] The invention provides nucleotide analogs having the formula NTP-linker-inhibitor for synthesis of polynucleotides in an aqueous environment. With respect to the analogs of the form NTP-linker-inhibitor, NTP can be any nucleotide triphosphate, such as adenosine triphosphate (ATP), guanosine triphosphate (GTP), cytidine triphosphate (CTP), thymidine triphosphate (TTP), uridine triphosphate (UTP), nucleotide triphosphates, deoxyadenosine triphosphate (dATP), deoxyguanosine triphosphate (dGTP), deoxycytidine triphosphate (dCTP), deoxythymidine triphosphate (dTTP), or deoxyuridine triphosphate (dUTP).

[0044] The linker can be any molecular moiety that links the inhibitor to the NTP and can be cleaved. For example, the linkers can be cleaved by adjusting the pH of the surrounding environment. The linkers may also be cleaved by an enzyme that is activated at a given temperature, but inactivated at another temperature. In some embodiments, the linkers include disulfide bonds.

[0045] Linkers may, for example, include photocleavable, nucleophilic, or electrophilic cleavage sites. Photocleavable linkers, wherein cleavage is activated by a particular wavelength of light, may include benzoin, nitroveratryl, phenacyl, pivaloyl, sisyl, 2-hydroxy-cinamyl, coumarin-4-yl-methyl, or 2-nitrobenzyl based linkers.

[0046] Examples of nucleophilic cleavage sites include fluoride ion cleavable silicon-oxygen bonds or esters which may be cleaved in a basic solution. Electrophilically cleaved linkers may include acid induced cleavage sites which may comprise trityl, tert-butyloxycarbonyl groups, acetal groups, and p-alkoxybenzyl esters and amides. In certain aspects, a cleavable linker may include a cysteine residue as shown in FIG. 15.

[0047] The linker can be attached, for example, at the N4 of cytosine, the N3 or O4 of thymine, the N2 or N3 of guanine, and the N6 of adenine, or the N3 or O4 of uracil because attachment at a carbon results in the presence of a residual scar after removal of the polymerase-inhibiting group. The linker is typically on the order of at least about 10 Angstroms long, e.g., at least about 20 Angstroms long, e.g., at least about 25 Angstroms long, thus allowing the inhibitor to be far enough from the pyridine or pyrimidine to allow the enzyme to bind the NTP to the polynucleotide chain via the attached sugar backbone. In some embodiments, the cleavable linkers are self-cyclizing in that they form a ring molecule that is particularly non-reactive toward the growing nucleotide chain.

[0048] In certain aspects, a cleavable linker may include a variable number of methylene bridges on the NTP or the inhibitor side of a disulfide bond, including, for example, 1, 2, 3, or 4 methylene bridges as shown in FIGS. 14 and 16A-C. These methylene bridges may be used to increase the space between the NTP and the inhibitor. As noted above, the length of the cleavable linker may be selected in order to prevent the

inhibitor from interfering with coupling of the NTP to the synthesized polynucleotide. In some embodiments of the invention, the distance of the charged group to the NTP plays an important role in the effectiveness of inhibiting a subsequent nucleotide incorporation.

[0049] For example, in some embodiments using a charged moiety as an inhibitor, the charged moiety may be from about 5 to about 60 bonds away from the NTP. In some other embodiments, the charged moiety of the inhibitor may be from about 10 to about 40 bonds away from the NTP. In some other embodiments, the charged moiety of the inhibitor can be from about 10 to about 35 bonds away from the NTP. In some other embodiments, the charged moiety of the inhibitor may be from about 10 to about 30 bonds away from the NTP. In some other embodiments, the charged moiety of the inhibitor is from about 10 to about 20 bonds away from the NTP. The number of bonds between the charged moiety and the NTP may be increased by including additional methylene bridges.

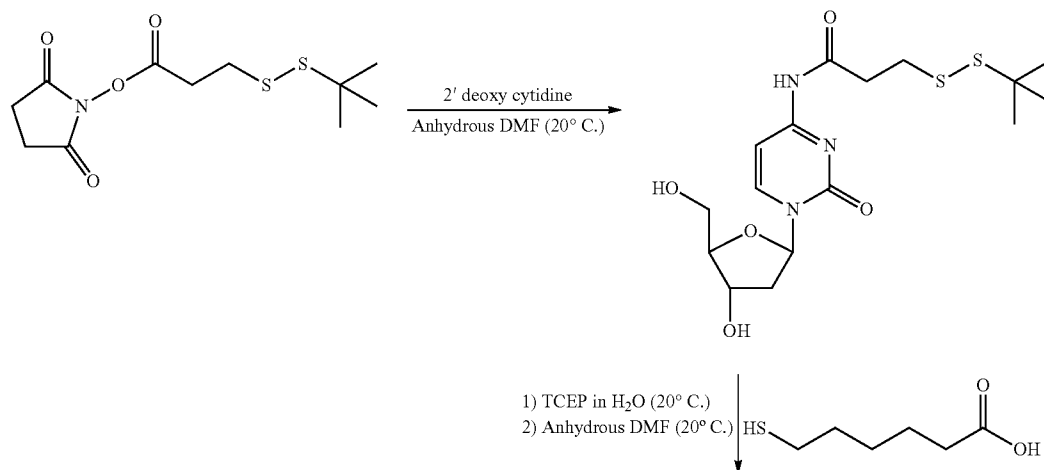
[0050] The nucleotide analogs can include any moiety linked to the NTP that inhibits the coupling of subsequent nucleotides by the enzyme. The inhibitory group can be a charged group, such as a charged amino acid, or the inhibitory group can be a group that becomes charged depending upon the ambient conditions. In some embodiments, the inhibitor may include a moiety that is negatively charged or capable of becoming a negatively charged. For example, an inhibitor may include a chain of phosphate groups (e.g., 1, 2, or 3, phosphates) as shown in FIGS. 16A-C, wherein additional phosphates increase the overall anionic charge of the inhibitor. In other embodiments, the inhibitor group is positively charged or capable of becoming positively charged. In some other embodiments, the inhibitor is an amino acid or an amino acid analog. The inhibitor may be a peptide of 2 to 20 units of amino acids or analogs, a peptide of 2 to 10 units of amino

acids or analogs, a peptide of 3 to 7 units of amino acids or analogs, a peptide of 3 to 5 units of amino acids or analogs. In some embodiments, the inhibitor includes a group selected from the group consisting of Glu, Asp, Arg, His, and Lys, and a combination thereof (e.g., Arg, Arg-Arg, Asp, Asp-Asp, Asp, Glu, Glu-Glu, Asp-Glu-Asp, Asp-Asp-Glu or AspAspAspAsp, etc.). Peptides or groups may be combinations of the same or different amino acids or analogs. In certain embodiments, a peptide inhibitor may be acetylated to discourage errant bonding of free amino groups. The inhibitory group may also include a group that reacts with residues in the active site of the enzyme thus interfering with the coupling of subsequent nucleotides by the enzyme. The inhibitor may have a charged group selected from the group consisting of $-\text{COO}^-$, $-\text{NO}_2$, $-\text{PO}_4^-$, $-\text{PO}_3^-$, $-\text{SO}_2^-$, or $-\text{NR}_3^+$ where each R may be H or an alkyl group. In other embodiments, the inhibitor moiety does not comprise a $-\text{PO}_4$ group.

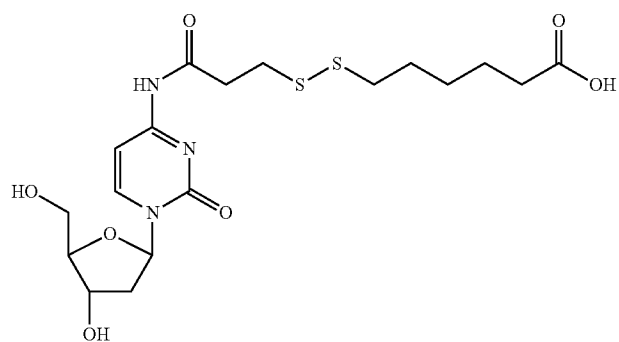
[0051] An example of a nucleotide analog of the type NTP-linker-inhibitor is shown in FIG. 1A. The analog in FIG. 1A includes an inhibitory (-Asp-Asp-) group linked to the N4 position of dCTP through a disulfide ($-\text{S}-\text{S}-$) bond while providing an unblocked, unmodified 3'-OH on the sugar ring. The linker is constructed such that all linker atoms (including the 2nd incorporation-inhibiting moiety) can be removed, thereby allowing the nascent DNA strand to revert to natural nucleotides. As shown in FIG. 1B, an aqueous reducing agent, such as tris(2-carboxyethyl) phosphine (TCEP) or dithiothreitol (DTT), can be used to cleave the $-\text{S}-\text{S}-$ bond, resulting in the loss of the inhibitor function (deblocking). As shown in FIG. 1B, a self-cyclizing linker can be incorporated, resulting in a cyclic oxidized tetrahydrothiophene leaving group that is easily removed from the reagent solution at the conclusion of nucleotide synthesis.

[0052] Exemplary schemes for synthesizing dCTP analogs of FIG. 1A are shown below in Schemes 1A and 1B:

Scheme 1A

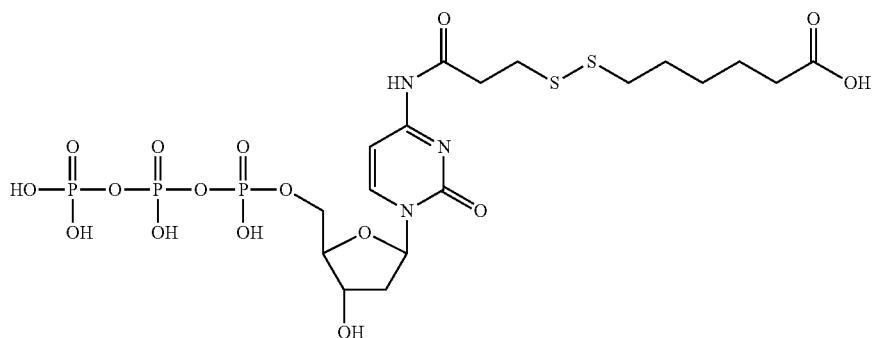


-continued

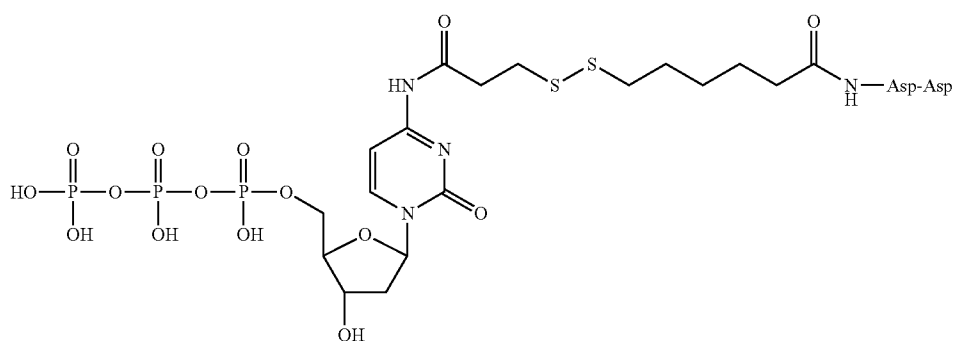


1) POCl_3 (20° C.)
2) $\text{PO}(\text{OCH}_3)_3$ (-15° C. - 0° C.)

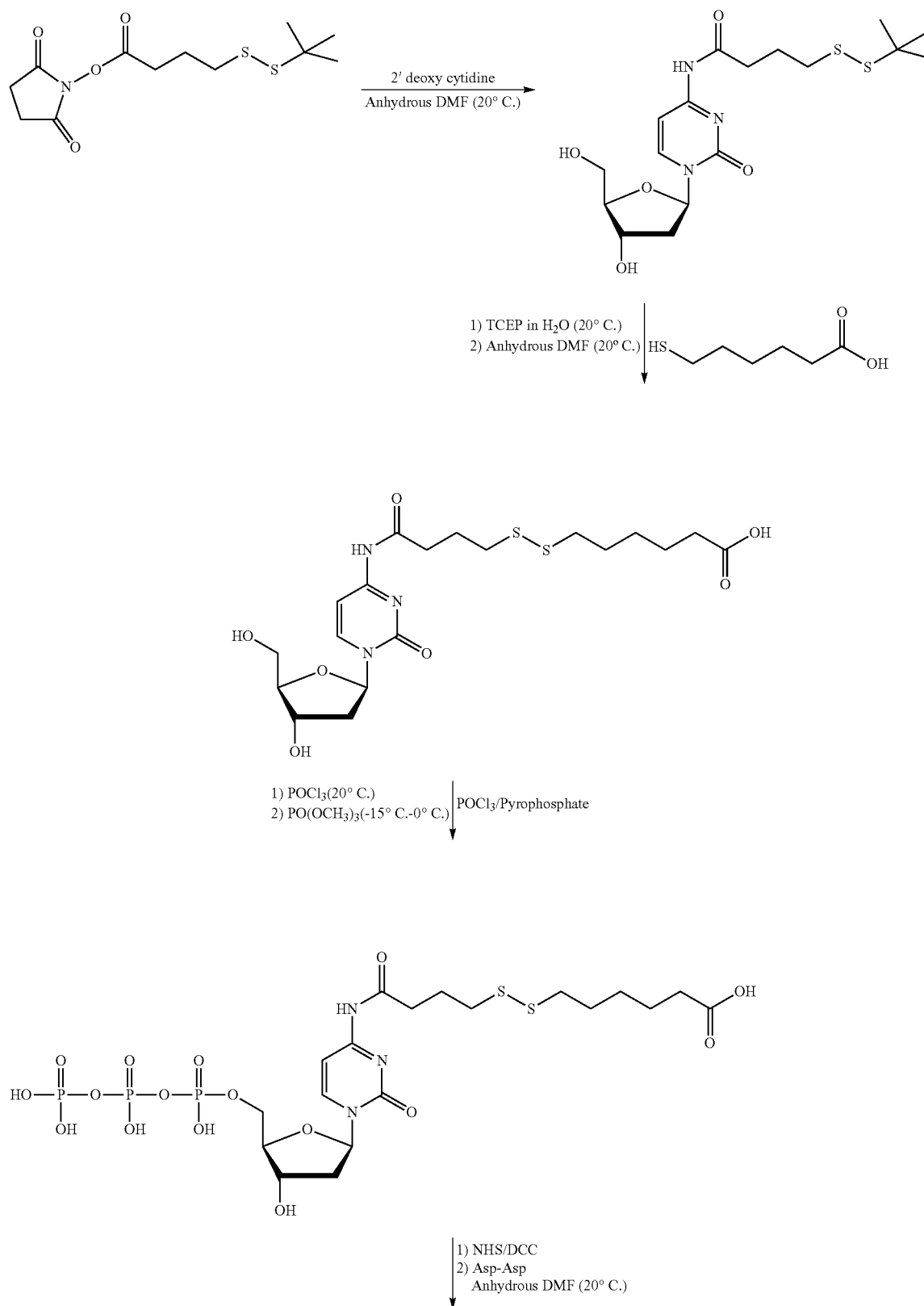
POCl_3 /Pyrophosphate

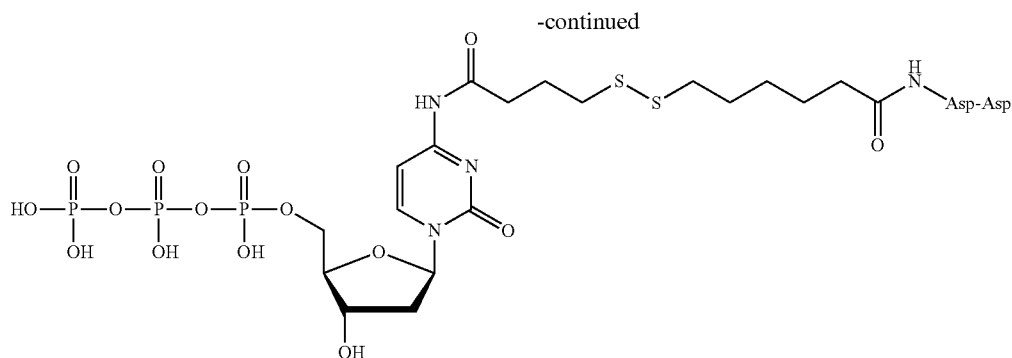


1) NHS/DCC
2) Asp-Asp
Anhydrous DMF (20° C.)



Scheme 1B

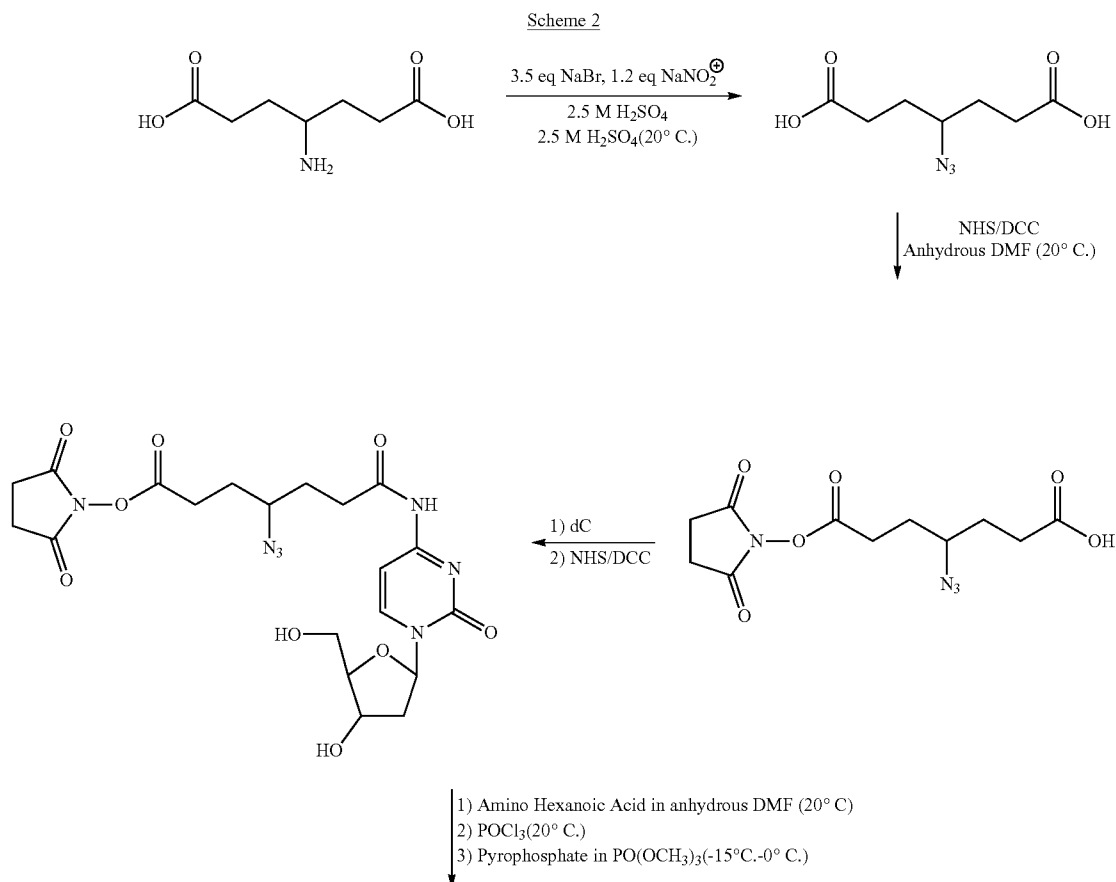


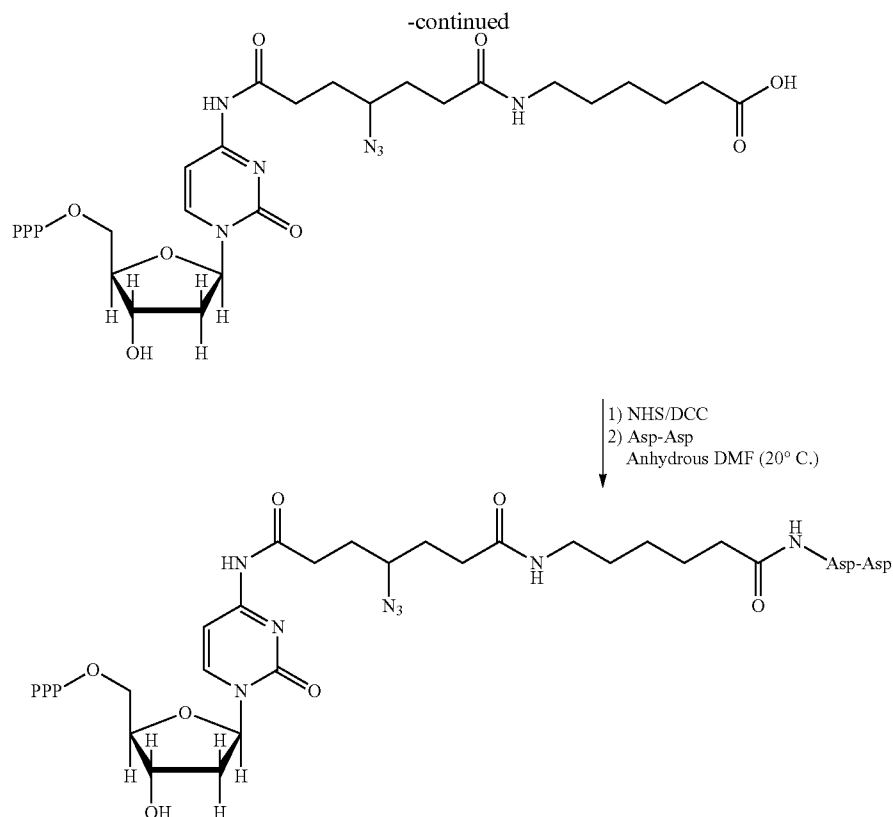


[0053] In a fashion analogous to Schemes 1A and 1B, nucleotide analogs of the type NTP-linker-inhibitor can also be formed by attaching the linker-inhibitor moiety to the N6 of adenine (FIG. 2), the N2 of guanine (FIG. 3), the N3 of thymine (FIG. 4), or the N3 of uracil (FIG. 5), thereby providing analogs of the “naturally-occurring” dNTPs, as well as a deoxyuracil nucleotide (dUTP). While it is unlikely that

there will be wide use of a dUTP, the synthesis is straightforward based upon the chemistry.

[0054] The invention is not limited to the linking chemistry of Schemes 1A and 1B, however, as carbamate, amide, or other self-eliminating linkages could also be employed. For example, nucleotides can also be prepared with Staudinger linkers, as shown in Scheme 2.





[0055] A deoxycytidine triphosphate (dCTP) analog created with a Staudinger linker (Scheme 2) to an Asp-Asp blocking group is shown in FIG. 6. As shown in FIG. 6, the Staudinger dCTP analog undergoes cleavage under aqueous conditions with the addition of azide and triphenylphosphine. The Staudinger analog shown in FIG. 6 is also suitable for nucleotide extension using nucleotidyl transferases, such as TdT, as described above and exemplified in FIGS. 1-5. While not shown explicitly in the FIGS., one of skill in the art can use Scheme 2 in conjunction with the suitable reactant to produce other nucleotide analogs having Staudinger linkers as needed for complete de novo nucleotide synthesis. In a fashion analogous to FIG. 6, nucleotide analogs of Scheme 2 can be formed by attaching the Staudinger moiety to the N6 of adenine, the N2 of guanine, the N3 of thymine, or the N3 of uracil, thereby providing analogs of the “naturally-occurring” dNTPs, as well as a deoxyuracil nucleotide (dUTP).

[0056] The methodologies of Scheme 1A can be used to produce corresponding ribonucleotide analogs, e.g., as shown in FIGS. 7-10, by starting with the appropriate ribonucleotide reactants. Ribonucleotide analogs comprising the Staudinger linker can also be created using Scheme 2 in order to form the needed ribonucleotide analogs, including, e.g., CTP analogs, as shown in FIG. 12. Furthermore, all of the ribonucleotide analogs, i.e., C, A, T, G, U, can be formed using a reaction similar to Scheme 2.

Enzymes

[0057] The methods of the invention employ nucleotidyl transferases to assemble the nucleotide analogs into poly-

nucleotides. Nucleotidyl transferases include several families of related transferase and polymerase enzymes. Some nucleotidyl transferases polymerize deoxyribonucleotides more efficiently than ribonucleotides, some nucleotidyl transferases polymerize ribonucleotides more efficiently than deoxyribonucleotides, and some nucleotidyl transferases polymerize ribonucleotides and deoxyribonucleotides at approximately the same rate.

[0058] Of particular import to the invention, transferases having polymerase activity, such as terminal deoxynucleotidyl transferase (TdT), are capable of catalyzing the addition of deoxyribonucleotides to the 3' end of a nucleotide chain, thereby increasing chain length in DNA nucleotides. TdT will only catalyze the addition of 1-2 ribonucleotides to the growing end of a DNA strand which could be useful in the construction of site specific DNA-RNA chimeric polynucleotides. In particular, calf thymus TdT, sourced from engineered *E. coli*, is suitable for use with the invention and available from commercial sources such as Thermo Scientific (Pittsburgh, Pa.). The amino acid sequence corresponding to calf TdT is listed in Table 1 as SEQ ID NO. 1.

TABLE 1

Amino Acid Sequence of Bovine TdT	
SEQ ID NO. 1:	
MAQQRQHQL	PMDPLCTASS GPRKKRPRQV GASMASPPHD
IKFQNLVLF	LEKKMGTTTR NFLMELARRK GFRVENELSD
SVTHIVAENN	SGSEVLEWLQ VQNIKASSQL ELLDVSWLIE

TABLE 1-continued

Amino Acid Sequence of Bovine TdT			
SMGAGKPVEI	TGKHQLVVRT	DYSATPNPGF	QKTPPLAVKK
ISQYACQRKT	TLNNYNHIFT	DAFEILAENS	EFKENEVSIV
TFMRAASVLK	SLPFTIISMK	DTEGIPCLGD	KVKCIIEEII
EDGESSEVKA	VLNDERYQSF	KLFTSVFVG	LKTSEKWFRM
GFRSLSKIMS	DKTLKFTKMQ	KAGFLYYEDL	VSCVTRAEAE
AVGVLVKEAV	WAFLPDAFVT	MTGGFRRGKK	IGHDVFLLIT
SPGSAEDEEQ	LLPKVINLWE	KKGLLLYYDL	VESTFEKFKL
PSRQVDTLDH	FQKCFLLIKL	HHQRVDSSKS	NQOEGKTWKA
IRVDLVMCPY	ENRAFALLGW	TGSRQFERDI	RRYATHERKM
MLDNHALYDK	TKRVFLKAES	EEEIFAHLGL	DYIEPWERN

The nucleotide sequence corresponding to calf TdT is listed in Table 2 as SEQ ID NO. 2.

TABLE 2

Nucleic Acid Sequence of Bovine TdT	
SEQ ID NO. 2:	
ctcttctgga	gataccactt gatggcacag cagaggcagc
atcagcgtct	tcccatggat ccgctgtgca cagcctcctc
aggccctcgg	aagaagagac ccaggcaggt ggggtgcctca
atggcctccc	ctcctcatga catcaagttt caaaatttgg
tctcttctat	tttgagaag aaaatgggaa ccaccgcag
aaacttcctc	atggagctgg ctgcaaggaa aggtttcagg
gttgaaaatg	agctcagtga ttctgtcacc cacattgtag
cagaaaacaa	ctctggttca gaggttctcg agtggttca
ggtacagaac	ataagagcca gctcgcagct agaactcctt
gatgtctcct	ggctgatcga aagtatggga gcaggaaaac
cagtggagat	tacaggaaaa caccagcttg ttgtgagaac
agactattca	gtaccccaaa acccaggctt ccagaagact
ccaccacttg	ctgtaaaaaa gatctcccag tacgcgtgtc
aaagaaaaac	cactttgaac aactataacc acatattcac
ggatgccttt	gagatactgg ctgaaaattc tgagttaaaa
gaaaatgaag	tctcttatgt gacatttatg agagcagctt
ctgtacttaa	atctctgccca ttcacaatca tcagtatgaa
ggatacagaa	ggaattccct gcctggggga caaggtgaag
tgtatcatag	aggaaattat tgaagatgga gaaagttctg
aagttaaagc	tgtgttaaat gatgaacgat atcagtcctt
caaactcttt	acttctgttt ttggagtgga actgaagaca
tctgagaaat	gggtcaggat ggggttcaga tctctgagta

TABLE 2-continued

Nucleic Acid Sequence of Bovine TdT			
aaataatgtc	agacaaaacc	ctgaaattca	caaaaatgca
gaaagcagga	tttctctatt	atgaagacct	tgctcagctgc
gtgaccaggg	ccgaagcaga	ggcggttggc	gtgctgggtta
aagaggctgt	gtgggcattt	ctgccggatg	cctttgtcac
catgacagga	ggattccgca	ggggttaagaa	gattgggcat
gatgtagatt	ttttaattac	cagcccagga	tcagcagagg
atgaagagca	acttttgcct	aaagtataaa	acttatggga
aaaaaaggga	ttacttttat	attatgacct	tgtggagtca
acatttgaaa	agttcaagtt	gccaagcagg	caggtggata
ctttagatca	ttttcaaaaa	tgctttctga	ttttaaaatt
gcaccatcag	agagtagaca	gtagcaagtc	caaccagcag
gaaggaaga	cctggaagc	catccgtgtg	gacctgggtta
tgtgccccta	cgagaaccgt	gcctttggcc	tgctaggtctg
gactggctcc	cggcagtttg	agagagacat	ccggcgctat
gccacacacg	agcgaagat	gatgctggat	aaccacgctt
tatatgacaa	gaccaagagg	gtattttctca	aagcggaaaag
tgaagaagaa	atctttgcac	atctgggatt	ggactacatt
gaacatggg	aaagaaatgc	ttaggagaaa	gctgtcaact
tttttctttt	ctgttctttt	tttcagggtta	gacaaattat
gcttcatatt	ataatgaaag	atgccttagt	caagtttggg
attctttaca	ttttaccaag	atgtagattg	cttctagaaa
taagtagttt	tggaaacgtg	atcaggcacc	ccctgggtta
tgctctggca	agccatttgc	aggactgatg	tgtagaactc
gcaatgcatt	ttccatagaa	acagtgttgg	aattgggtggc
tcatttcocag	ggaagtccat	caaagccac	tttgcccaca
gtgtagctga	aatactgtat	acttgccaat	aaaaatagga
aac			

[0059] While commercially-available TdT is suitable for use with the methods of the invention, modified TdT, e.g., having an amino acid sequence at least 95% in common with SEQ ID NO. 1, e.g., having an amino acid sequence at least 98% in common with SEQ ID NO. 1, e.g., having an amino acid sequence at least 99% in common with SEQ ID NO. 1, may be used with the methods of the invention. An organism that expresses a suitable nucleotidyl transferase may comprise a nucleic acid sequence at least 95% in common with SEQ ID NO. 2, e.g., at least 98% in common with SEQ ID NO. 2, e.g., at least 99% in common with SEQ ID NO. 2. In some instances, a modified TdT will result in more efficient generation of polynucleotides, or allow better control of chain length. Other modifications to the TdT may change the release characteristics of the enzyme, thereby reducing the need for aqueous reducing agents such as TCEP or DTT.

[0060] For the synthesis of RNA polynucleotides, a nucleotidyl transferase like *E. coli* poly(A) polymerase can be used to catalyze the addition of ribonucleotides to the 3' end of a ribonucleotide initiator. In other embodiments, *E. coli* poly(U) polymerase may be more suitable for use with the methods of the invention. Both *E. coli* poly(A) polymerase and *E. coli* poly(U) polymerase are available from New England Biolabs (Ipswich, Mass.). These enzymes may be used with 3'unblocked reversible terminator ribonucleotide triphosphates (rNTPs) to synthesize RNA. In certain embodiments, RNA may be synthesized using 3'blocked, 2'blocked, or 2'-3'blocked rNTPs and poly(U) polymerase or poly(A) polymerase. The amino acid and nucleotide sequences for *E. coli* Poly(A) polymerase and *E. coli* Poly(U) polymerase are reproduced below. Modified *E. coli* Poly(A) polymerase or *E. coli* Poly(U) polymerase may be suitable for use with the methods of the invention. For example, an enzyme, having an amino acid sequence at least 95% in common with SEQ ID NO. 3, e.g., having an amino acid sequence at least 98% in common with SEQ ID NO. 3, e.g., having an amino acid sequence at least 99% in common with SEQ ID NO. 3, may be used with the methods of the invention. An organism that expresses a suitable enzyme may comprise a nucleic acid sequence at least 95% in common with SEQ ID NO. 4, e.g., at least 98% in common with SEQ ID NO. 4, e.g., at least 99% in common with SEQ ID NO. 4. Alternatively, an enzyme having an amino acid sequence at least 95% in common with SEQ ID NO. 5, e.g., having an amino acid sequence at least 98% in common with SEQ ID NO. 5, e.g., having an amino acid sequence at least 99% in common with SEQ ID NO. 5, may be used with the methods of the invention. An organism that expresses a suitable enzyme may comprise a nucleic acid sequence at least 95% in common with SEQ ID NO. 6, e.g., at least 98% in common with SEQ ID NO. 6, e.g., at least 99% in common with SEQ ID NO. 6.

TABLE 3

Amino Acid Sequence of <i>E. coli</i> Poly(A) polymerase
SEQ ID NO. 3:
MFTRVANFCR KVLRSRESEA EQAVARPOVT VIPREQHAIS
RKDISENALK VMYRLNKAGY EAWLVGGGVR DLLLGKKPKD
FDVTTNATPE QVRKLFNRCR LVGRRFRLAH VMFGPEIIEV
ATFRGHHEGN VSDRTTSQRG QNGMLLRDNI FGSIEEDAQR
RDFTINSLYY SVADFTVRDY VGGMKDLKDG VIRLIGNPET
RYREDPVRL RAVRFAAKLG MRISPETAEP IPRLATLLND
IPPARLFEEES LKLLQAGYGY ETYKLLCEYH LFQPLFPPTIT
RYFTENGDSP MERIEQVLK NTDTRIHNND RVNPAFLFAA
MFWYPLLETA QKIAQESGLT YHDAFALAMN DVLDEACRSL
AIPKRLTTLT RDIWQLQLRM SRRQ GKRAWK LLEHPKFRAA
YDLLALRAEV ERNAELQRLV KWWGEFQVSA PPDQKGLMNE
LDEEPSPRRR TRRPRKRAPR REGTA

TABLE 4

Nucleotide Sequence of <i>E. coli</i> Poly(A) polymerase
SEQ ID NO. 4:
atttttaccg gagtcgctaa tttttgccgc aaggtgctaa
gccgcgagga aagcgaggct gaacaggcag tccccgtcc
acaggtgacg gtgatccccg gtgagcagca tgctatttcc
cgcaaagata tcagtgaaaa tgcctgaag gtaatgtaca
ggctcaataa agcgggatac gaagcctggc tggttggcgg
cggcgtgcgc gacctgttac ttggcaaaaa gccgaaagat
tttgacgtaa ccactaacgc cagcctgag caggtgcgca
aactgttccg taactgccgc ctggtgggtc gccgtttccg
tctggctcat gtaatgtttg gcccgagat tatcgaagtt
gcgaccttcc gtggacacca cgaaggtaac gtcagcgacc
gcacgacctc ccaacgcggg caaacggca tgttgctgcg
cgacaacatt ttcggtcca tcgaagaaga cgcccagcgc
cgcgatttca ctatcaacag cctgtattac agcgtagcgg
attttaccgt ccgtgattac gttggcgcca tgaaggatct
gaaggacggc gttatccgtc tgattggtaa cccggaaacg
cgctaccgtg aagatccggt acgtatgctg cgcgcggtac
gttttgccgc caaattgggt atgcgcatca gcccgaaac
cgcgaaccg atccctcgcc tcgctacctt gctgaacgat
atcccaccgg cagcctgtt tgaagaatcg cttaaactgc
tacaagcggg ctacggttac gaaacctata agctgtgttg
tgaatatcat ctgttccagc cgctgttccc gaccattacc
cgctacttca cggaaatgg cgacagcccg atggagcgga
tcattgaaca ggtgctgaag aataccgata cgcgtatcca
taacgatatg cgcgtgaacc cggcgttccct gtttgccgcc
atgttctggt acccactgct ggagacggca cagaagatcg
cccagaaaag cggcctgacc tatcacgacg ctttcgcgct
ggcgatgaac gacgtgctgg acgaagcctg ccgttcactg
gcaatcccg aacgtctgac gacattaacc cgcgatatct
ggcagttgca gttgcgtatg tcccgtcgtc agggtaaacg
cgcagtgaaa ctgctggagc atcctaagtt ccgtgcggct
tatgacctgt tggccttgcg agctgaagtt gagcgtaacg
ctgaactgca gcgtctggtg aaatgggtggg gtgagttcca
ggtttccgcg ccaccagacc aaaaagggat gctcaacgag
ctggatgaag aaccgtcacc gcgtcgctcg actcgctcgc
cacgcaaacg cgcaccacgt cgtgagggta ccgcatga

The nucleotide sequence corresponding to *E. coli* poly(A) polymerase is listed in Table 4 as SEQ ID NO. 4.

TABLE 5

Amino Acid Sequence of <i>E. coli</i> Poly(U) polymerase			
SEQ ID NO. 5:			
GSHMSYQKVP	NSHKEFTKFC	YEVYNEIKIS	DKEFKEKRAA
LDTLRLCLKR	ISPDAELVAF	GSLESGALK	NSDMDLCVLM
DSRVQSDTIA	LQFYEEILAE	GFEGKFLQRA	RIPPIKLTSD
TKNGFGASFQ	CDIGFNNRLA	IHNLTLLSSY	TKLDARLKPM
VLLVKHWAKR	KQINSPYFGT	LSSYGYVLMV	LYYLIHVIKP
PVFPNLLLSR	LKQEKIVDGF	DVGFDKLED	IPPSQNYSSL
GSLHGFRRF	YAYKFEPRK	VVTFRRPDGY	LTKQEKGWTS
ATEHTGSADQ	IKDRYILAI	EDPFEISHNV	GRTVSSSGLY
RIRGEFMAAS	RLNRSRYPI	PYDSLFEAA	

The nucleotide sequence corresponding to *E. coli* poly(U) polymerase is listed in Table 6 as SEQ ID NO. 6.

TABLE 6

Nucleotide Sequence of <i>E. coli</i> Poly(A) polymerase			
SEQ ID NO. 6:			
ggcagccata	tgaactatca	gaaagtgcgc	aacagccata
aagaatttac	caaatattgc	tatgaagtgt	ataacgaaat
taaaattagc	gataaagaat	ttaaagaaaa	acgcgcggcg
ctggataccc	tgcgcctgtg	cctgaaacgc	attagcccg
atgcggaaat	gggtggcgtt	ggcagcctgg	aaagcggcct
ggcgtgaaa	aacagcgata	tggatctgtg	cgtgctgatg
gatagcgcgc	tgcagagcga	taccattgcg	ctgcagtttt
atgaagaact	gattgcggaa	ggctttgaag	gcaaatctct
gcagcgcgcg	cgcattccga	ttattaaact	gaccagcgat
acaaaaaacg	gctttggcgc	gagctttcag	tgcgatattg
gctttaacaa	ccgcctggcg	attcataaca	ccctgctgct
gagcagctat	acaaaactgg	atgcgcgcct	gaaaccgatg
gtgctgctgg	tgaacatttg	ggcgaaacgc	aaacagatta
acagcccgta	ttttggcacc	ctgagcagct	attggctatgt
gctgatggcg	ctgtattatc	tgattcatgt	gattaaaccg
ccggtgtttc	cgaacctgct	gctgagcccg	ctgaaacagg
aaaaaattgt	ggatggcgtt	gatgtgggct	ttgatgataa
actggaagat	attccgccga	gccagaacta	tagcagcctg
ggcagcctgc	tgcattggctt	ttttcgcttt	tatgcgtata
aatttgaacc	gcgcgaaaaa	gtggtgacct	ttcgccgcc
ggatggctat	ctgaccaaac	aggaaaaagg	ctggaccagc
gcgaccgaac	ataccggcag	cgcggatcag	attattaaag
atcgctatat	tctggcgatt	gaagatccgt	ttgaaattag
ccataacgtg	ggccgcaccg	tgagcagcag	cggcctgtat
cgcattcgcg	gcgaatttat	ggcggcgagc	cgccctgctga
acagcccgag	ctatccgatt	ccgtatgata	gcctgtttga
agaagcg			

[0061] As discussed above, the inhibitor coupled to the nucleotide analog will cause the transferase, e.g., TdT, to not release from the polynucleotide or prevent other analogs from being incorporated into the growing chain. A charged moiety results in better inhibition, however, research suggests that the specific chemical nature of the inhibitor is not particularly important. For example, both phosphates and acidic peptides can be used to inhibit enzymatic activity. See, e.g., Bowers et al., *Nature Methods*, vol. 6, (2009) p. 593-95, and U.S. Pat. No. 8,071,755, both of which are incorporated herein by reference in their entireties. In some embodiments, the inhibitor will include single amino acids or dipeptides, like -(Asp)₂, however the size and charge on the moiety can be adjusted, as needed, based upon experimentally determined rates of first nucleotide incorporation and second nucleotide incorporation. That is, other embodiments may use more or different charged amino acids or other biocompatible charged molecule.

[0062] Other methods of nucleotide synthesis may be used to build de novo oligonucleotides in a template independent fashion using nucleotidyl transferases or modified nucleoti-

dyl transferases. In one embodiment, the polymerase/transferase enzymes can be modified so that they cease nucleotide addition when they encounter a modification to the phosphate of a 3'-unmodified dNTP analog. This scheme would require a deblocking reagent/reaction that modifies the phosphate end of the nucleotide analog, which frees up the nascent strand for subsequent nucleotide incorporation. Preferred embodiments of this approach would use nucleotide analogs modified only at the phosphates (alpha, beta or gamma) although modifications of the purine/pyrimidine base of the nucleotide are allowed.

[0063] Another embodiment for using non-template dependent polymerase/transferase enzymes would be to use protein engineering or protein evolution to modify the enzyme to remain tightly bound and inactive to the nascent strand after each single nucleotide incorporation, thus preventing any subsequent incorporation until such time as the polymerase/transferase is released from the strand by use of a releasing reagent/condition. Such modifications would be selected to allow the use of natural unmodified dNTPs instead of reversible terminator dNTPs. Releasing reagents could be high salt buffers, denaturants, etc. Releasing conditions could be high temperature, agitation, etc. For instance, mutations to the Loop1 and SD1 regions of TdT have been shown to dramatically alter the activity from a template-independent activity to more of a template dependent activity. Specific mutations of interest include but are not limited to $\Delta_{384/391/392}$, del loop1 (386→398), L398A, D339A, F401A, and Q402K403C404→E402R403S404. Other means of accomplishing the goal of a post-incorporation tight binding (i.e., single turnover) TdT enzyme could include mutations to the residues responsible for binding the three phosphates of the initiator strand including but not limited to K261, R432, and R454.

[0064] Another embodiment for using non-template dependent polymerase/transferase enzymes would be to use protein engineering or protein evolution to modify the enzyme to accept 3'-blocked reversible terminators with high efficiency. Most naturally occurring polymerase/transferase enzymes will not incorporate 3'-blocked reversible terminators due to steric constraints in the active site of the enzyme. Modifying either single or several aa residues in the active site of the enzyme can allow the highly efficient incorporation of 3'-blocked reversible terminators into a support bound initiator in a process completely analogous to that described above. After incorporation, the 3'-reversible terminator is removed with a deblocking reagent/condition thus generating a completely natural (scarless) single strand molecule ready for subsequent controlled extension reactions. There are residues close to the 3'-OH of the incoming dNTP which explains the propensity of TdT for incorporating ribonucleotide triphosphates as readily as deoxyribonucleotide triphosphates; residues including but not limited to those between $\beta 1$ and $\beta 2$ especially R334, Loop1, and those between $\alpha 13$ and $\alpha 14$, especially R454, are likely targets for mutagenesis to accommodate the bulk of 3'-reversible terminator groups and allow their efficient incorporation. In certain embodiments additional amino acid changes may be required to compensate for these residue alterations made to accommodate a 3'-reversible terminator. Another embodiment for using template-dependent polymerases would be to use the either 3'blocked or 3'unblocked dNTP analogs with a plurality of primer-template pairs attached to a solid support where the template is a

nucleic acid analog that supports polymerase mediated primer extension of any of the four bases as specified by the user.

[0065] Another embodiment for using non-template dependent polymerase/transferase enzymes can use protein engineering or protein evolution to modify the enzyme to optimize the use of each of the four different nucleotides or even different modified nucleotide analogs in an analog specific manner. Nucleotide specific or nucleotide analog specific enzyme variants could be engineered to possess desirable biochemical attributes like reduced K_m or enhanced addition rate which would further reduce the cost of the synthesis of desired polynucleotides.

Solid State Synthesis

[0066] The methods of the invention can be practiced under a variety of reaction conditions, however the orderly construction and recovery of desired polynucleotides will, in most cases, require a solid support to which the polynucleotides can be grown. In some embodiments, the methods include the enzymatically-mediated synthesis of polynucleotides on a solid support, as illustrated in FIG. 7. When used in conjunction with the NTP, linker, inhibitor analogs discussed above, it is possible to construct specific polynucleotide sequences of DNA as well as RNA by using, for example, TdT or poly(A) polymerase in an aqueous environment. As shown in FIG. 13, the TdT can be used to effect the stepwise construction of custom polynucleotides by extending the polynucleotide sequence a stepwise fashion. As discussed previously, the inhibitor group of each NTP analog causes the enzyme to stop with the addition of a nucleotide. After each nucleotide extension step, the reactants are washed away from the solid support prior to the removal of the inhibitor by cleaving the linker, and then new reactants are added, allowing the cycle to start anew. At the conclusion of n cycles of extension-remove-deblocking-wash, the finished full-length, single-strand polynucleotide is complete and can be cleaved from the solid support and recovered for subsequent use in applications such as DNA sequencing or PCR. Alternatively, the finished, full-length, single-strand polynucleotide can remain attached to the solid support for subsequent use in applications such as hybridization analysis, protein or DNA affinity capture. In other embodiments, partially double-stranded DNA can be used as an initiator, resulting in the synthesis of double-stranded polynucleotides.

[0067] Solid supports suitable for use with the methods of the invention may include glass and silica supports, including beads, slides, pegs, or wells. In some embodiments, the support may be tethered to another structure, such as a polymer well plate or pipette tip. In some embodiments, the solid support may have additional magnetic properties, thus allowing the support to be manipulated or removed from a location using magnets. In other embodiments, the solid support may be a silica coated polymer, thereby allowing the formation of a variety of structural shapes that lend themselves to automated processing.

Synthesizers

[0068] To capitalize on the efficiency of the disclosed methods, an aqueous phase DNA synthesizer can be constructed to produce desired polynucleotides in substantial quantities. In one embodiment, a synthesizer will include four wells of the described NTP analog reagents, i.e., dCTP, dATP, dGTP, and

dTTP, as well as TdT at concentrations sufficient to effect polynucleotide growth. A plurality of initiating sequences can be attached to a solid support that is designed to be repeatedly dipped into each of the four wells, e.g., using a laboratory robot. The robot could be additionally programmed to rinse the solid support in wash buffer between nucleotide additions, cleave the linking group by exposing the support to a deblocking agent, and wash the solid support a second time prior to moving the solid support to the well of the next desired nucleotide. With simple programming, it is possible to create useful amounts of desired nucleotide sequences in a matter of hours, and with substantial reductions hazardous waste. Ongoing synthesis under carefully controlled conditions will allow the synthesis of polynucleotides with lengths in the thousands of base pairs. Upon completion, the extension products are released from the solid support, whereupon they can be used as finished nucleotide sequences.

[0069] A highly parallel embodiment could consist of a series of initiator-solid supports on pegs in either 96 or 384 well formats that could be individually retracted or lowered so that the pegs can be indexed to contact the liquids in the wells in a controlled fashion. The synthesizer could thus consist of the randomly addressable peg device, four enzyme-dNTP analog reservoirs in the same format as the peg device (96 or 384 spacing), additional reagent reservoirs (washing, deblocking, etc.) in the same format as the peg device (96 or 384 spacing), and a transport mechanism (e.g., a laboratory robot) for moving the peg device from one reservoir to another in a user programmable controlled but random access fashion. Care must be taken to avoid contaminating each of the four enzyme-dNTP reservoirs since the contents are reused throughout the entire synthesis process to reduce the cost of each polynucleotide synthesis.

[0070] In alternative embodiments, the reagents (e.g., nucleotide analogs, enzymes, buffers) will be moved between solid supports, allowing the reagents to be recycled. For example a system of reservoirs and pumps can move four different nucleotide analog solutions, wash buffers, and/or reducing agent solutions between one or more reactors in which the oligonucleotides will be formed. The reactors and pumps can be conventional, or the devices may be constructed using microfluidics. Because of the non-anhydrous (aqueous) nature of the process, no special care needs to be taken in the design of the hardware used to eliminate exposure to water. The synthesis process can take place with only precautions to control evaporative loss. A highly parallel embodiment could consist of a monolithic series of initiator-solid supports on pegs in either 96 or 384 well format that can be interfaced to a series of wells in the same matching format. Each well would actually be a reaction chamber that is fed by four enzyme-dNTP analog reservoirs and additional reagent reservoirs (washing, deblocking, etc.) with appropriate valves. Provisions would be made in the fluidics logic to recover the enzyme-dNTP reactants in a pristine fashion after each extension reaction since they are reused throughout the entire synthesis process to reduce the cost of each polynucleotide synthesis. In other embodiments, a system of pipetting tips could be used to add and remove reagents.

[0071] In certain aspects, polynucleotides may be synthesized using microfluidic devices and/or inkjet printing technology. An exemplary microfluidic polynucleotide synthesis device is shown in FIG. 17 for illustrative purposes and not to scale. Microfluidic channels 255, including regulators 257, couple reservoirs 253 to a reaction chamber 251 and an outlet

channel 259, including a regulator 257 can evacuate waste from the reaction chamber 251. Microfluidic devices for polynucleotide synthesis may include, for example, channels 255, reservoirs 253, and/or regulators 257. Polynucleotide synthesis may occur in a microfluidic reaction chamber 251 which may include a number of anchored synthesized nucleotide initiators which may include beads or other substrates anchored or bound to an interior surface of the reaction chamber and capable of releasably bonding a NTP analog or polynucleotide initiator. The reaction chamber 251 may include at least one intake and one outlet channel 259 so that reagents may be added and removed to the reaction chamber 251. The microfluidic device may include a reservoir 253 for each respective NTP analog. Each of these NTP analog reservoirs 253 may also include an appropriate amount of TdT or any other enzyme which elongates DNA or RNA strands without template direction. Additional reservoirs 253 may contain reagents for linker/inhibitor cleavage and washing. These reservoirs 253 can be coupled to the reaction chamber 251 via separate channels 255 and reagent flow through each channel 255 into the reaction chamber 251 may be individually regulated through the use of gates, valves, pressure regulators, or other means. Flow out of the reaction chamber 251, through the outlet channel 259, may be similarly regulated.

[0072] In certain instances, reagents may be recycled, particularly the NTP analog-enzyme reagents. Reagents may be drawn back into their respective reservoirs 253 from the reaction chamber 251 via the same channels 255 through which they entered by inducing reverse flow using gates, valves, pressure regulators or other means. Alternatively, reagents may be returned from the reaction chamber 251 to their respective reservoirs 253 via independent return channels. The microfluidic device may include a controller capable of operating the gates, valves, pressure, or other regulators 257 described above.

[0073] An exemplary microfluidic polynucleotide synthesis reaction may include flowing a desired enzyme-NTP analog reagent into the reaction chamber 251; after a set amount of time, removing the enzyme-NTP analog reagent from the reaction chamber 251 via an outlet channel 259 or a return channel; flowing a wash reagent into the reaction chamber 251; removing the wash reagent from the reaction chamber 251 through an outlet channel 259; flowing a de-blocking or cleavage reagent into the reaction chamber 251; removing the de-blocking or cleavage reagent from the reaction chamber 251 via an outlet channel 259 or a return channel; flowing a wash reagent into the reaction chamber 251; removing the wash reagent from the reaction chamber 251 through an outlet channel 259; flowing the enzyme-NTP analog reagent including the next NTP in the desired sequence to be synthesized into the reaction chamber 251; and repeating until the desired polynucleotide has been synthesized. After the desired polynucleotide has been synthesized, it may be released from the reaction chamber anchor or substrate and collected via an outlet channel 259 or other means.

[0074] In certain aspects, reagents and compounds, including NTP analogs, TdT and/or other enzymes, and reagents for linker/inhibitor cleavage and/or washing may be deposited into a reaction chamber using inkjet printing technology or piezoelectric drop-on-demand (DOD) inkjet printing technology. Inkjet printing technology can be used to form droplets, which can be deposited, through the air, into a reaction chamber. Reagent droplets may have volumes in the picoliter to nanoliter scale. Droplets may be introduced using inkjet

printing technology at a variety of frequencies including 1 Hz, 10 Hz, 100 Hz, 1 kHz, 2 kHz, and 2.5 kHz. Various reagents may be stored in separate reservoirs within the inkjet printing device and the inkjet printing device may deliver droplets of various reagents to various discrete locations including, for example, different reaction chambers or wells within a chip. In certain embodiments, inkjet and microfluidic technologies may be combined wherein certain reagents and compounds are delivered to the reaction chamber via inkjet printing technology while others are delivered via microfluidic channels or tubes. An inkjet printing device may be controlled by a computing device comprising at least a non-transitory, tangible memory coupled to a processor. The computing device may be operable to receive input from an input device including, for example, a touch screen, mouse, or keyboard and to control when and where the inkjet printing device deposits a droplet of reagent, the reagent it deposits, and/or the amount of reagent deposited.

[0075] In certain instances, a desired polynucleotide sequence may be entered into the computing device through an input device wherein the computing device is operable to perform the necessary reactions to produce the desired polynucleotide sequence by sequentially depositing the appropriate NTP analog, enzyme, cleavage reagent, and washing reagent, in the appropriate order as described above.

[0076] After synthesis, the released extension products can be analyzed by high resolution PAGE to determine if the initiators have been extended by the anticipated number of bases compared to controls. A portion of the recovered synthetic DNA may also be sequenced to determine if the synthesized polynucleotides are of the anticipated sequence.

[0077] Because the synthesizers are relatively simple and do not require the toxic components needed for phosphoramidite synthesis, synthesizers of the invention will be widely accessible for research institutions, biotechs, and hospitals. Additionally, the ability to reuse/recycle reagents will reduce the waste produced and help reduce the costs of consumables. The inventors anticipate that the methods and systems will be useful in a number of applications, such as DNA sequencing, PCR, and synthetic biology.

INCORPORATION BY REFERENCE

[0078] References and citations to other documents, such as patents, patent applications, patent publications, journals, books, papers, web contents, have been made throughout this disclosure. All such documents are hereby incorporated herein by reference in their entirety for all purposes.

EQUIVALENTS

[0079] Various modifications of the invention and many further embodiments thereof, in addition to those shown and described herein, will become apparent to those skilled in the art from the full contents of this document, including references to the scientific and patent literature cited herein. The subject matter herein contains important information, exemplification and guidance that can be adapted to the practice of this invention in its various embodiments and equivalents thereof.

 SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 6

<210> SEQ ID NO 1

<211> LENGTH: 520

<212> TYPE: PRT

<213> ORGANISM: Bos taurus

<400> SEQUENCE: 1

```

Met Ala Gln Gln Arg Gln His Gln Arg Leu Pro Met Asp Pro Leu Cys
1          5          10          15

Thr Ala Ser Ser Gly Pro Arg Lys Lys Arg Pro Arg Gln Val Gly Ala
20        25        30

Ser Met Ala Ser Pro Pro His Asp Ile Lys Phe Gln Asn Leu Val Leu
35        40        45

Phe Ile Leu Glu Lys Lys Met Gly Thr Thr Arg Arg Asn Phe Leu Met
50        55        60

Glu Leu Ala Arg Arg Lys Gly Phe Arg Val Glu Asn Glu Leu Ser Asp
65        70        75        80

Ser Val Thr His Ile Val Ala Glu Asn Asn Ser Gly Ser Glu Val Leu
85        90        95

Glu Trp Leu Gln Val Gln Asn Ile Arg Ala Ser Ser Gln Leu Glu Leu
100       105       110

Leu Asp Val Ser Trp Leu Ile Glu Ser Met Gly Ala Gly Lys Pro Val
115       120       125

Glu Ile Thr Gly Lys His Gln Leu Val Val Arg Thr Asp Tyr Ser Ala
130       135       140

Thr Pro Asn Pro Gly Phe Gln Lys Thr Pro Pro Leu Ala Val Lys Lys
145       150       155       160

Ile Ser Gln Tyr Ala Cys Gln Arg Lys Thr Thr Leu Asn Asn Tyr Asn
165       170       175

His Ile Phe Thr Asp Ala Phe Glu Ile Leu Ala Glu Asn Ser Glu Phe
180       185       190

Lys Glu Asn Glu Val Ser Tyr Val Thr Phe Met Arg Ala Ala Ser Val
195       200       205

Leu Lys Ser Leu Pro Phe Thr Ile Ile Ser Met Lys Asp Thr Glu Gly
210       215       220

Ile Pro Cys Leu Gly Asp Lys Val Lys Cys Ile Ile Glu Glu Ile Ile
225       230       235       240

Glu Asp Gly Glu Ser Ser Glu Val Lys Ala Val Leu Asn Asp Glu Arg
245       250       255

Tyr Gln Ser Phe Lys Leu Phe Thr Ser Val Phe Gly Val Gly Leu Lys
260       265       270

Thr Ser Glu Lys Trp Phe Arg Met Gly Phe Arg Ser Leu Ser Lys Ile
275       280       285

Met Ser Asp Lys Thr Leu Lys Phe Thr Lys Met Gln Lys Ala Gly Phe
290       295       300

Leu Tyr Tyr Glu Asp Leu Val Ser Cys Val Thr Arg Ala Glu Ala Glu
305       310       315       320

Ala Val Gly Val Leu Val Lys Glu Ala Val Trp Ala Phe Leu Pro Asp
325       330       335

Ala Phe Val Thr Met Thr Gly Gly Phe Arg Arg Gly Lys Lys Ile Gly
340       345       350

```

-continued

His Asp Val Asp Phe Leu Ile Thr Ser Pro Gly Ser Ala Glu Asp Glu
 355 360 365
 Glu Gln Leu Leu Pro Lys Val Ile Asn Leu Trp Glu Lys Lys Gly Leu
 370 375 380
 Leu Leu Tyr Tyr Asp Leu Val Glu Ser Thr Phe Glu Lys Phe Lys Leu
 385 390 395 400
 Pro Ser Arg Gln Val Asp Thr Leu Asp His Phe Gln Lys Cys Phe Leu
 405 410 415
 Ile Leu Lys Leu His His Gln Arg Val Asp Ser Ser Lys Ser Asn Gln
 420 425 430
 Gln Glu Gly Lys Thr Trp Lys Ala Ile Arg Val Asp Leu Val Met Cys
 435 440 445
 Pro Tyr Glu Asn Arg Ala Phe Ala Leu Leu Gly Trp Thr Gly Ser Arg
 450 455 460
 Gln Phe Glu Arg Asp Ile Arg Arg Tyr Ala Thr His Glu Arg Lys Met
 465 470 475 480
 Met Leu Asp Asn His Ala Leu Tyr Asp Lys Thr Lys Arg Val Phe Leu
 485 490 495
 Lys Ala Glu Ser Glu Glu Glu Ile Phe Ala His Leu Gly Leu Asp Tyr
 500 505 510
 Ile Glu Pro Trp Glu Arg Asn Ala
 515 520

<210> SEQ ID NO 2

<211> LENGTH: 1923

<212> TYPE: DNA

<213> ORGANISM: Bos taurus

<400> SEQUENCE: 2

```

ctcttctgga gataccactt gatggcacag cagaggcagc atcagcgtct tcccatggat    60
ccgctgtgca cagcctcttc aggcctcctg aagaagagac ccaggcaggt ggggtgcctca    120
atggcctccc ctctcatga catcaagttt caaaatttgg tcctcttcat tttggagaag    180
aaaaatggaa ccacccgcag aaacttcttc atggagctgg ctggaaggaa aggtttcagg    240
gttgaaaatg agctcagtga ttctgtcacc cacattgtag cagaaaacaa ctctggttca    300
gaggttctcg agtggcttca ggtacagaac ataagagcca gctcgagct agaactcctt    360
gatgtctcct ggctgatcga aagtatggga gcaggaaaac cagtggagat tacaggaaaa    420
caccagcttg ttgtgagaac agactattca gctaccccaa acccaggctt ccagaagact    480
ccaccacttg ctgtaaaaaa gatctcccag tacgcgtgtc aaagaaaaac cactttgaac    540
aactataacc acatattcac ggatgccttt gagatactgg ctgaaaattc tgagtttaaa    600
gaaaatgaag tctcttatgt gacatttatg agagcagctt ctgtacttaa atctctgcca    660
ttcacaatca tcagtatgaa ggatacagaa ggaattccct gcctggggga caaggtgaag    720
tgtatcatag aggaaattat tgaagatgga gaaagttctg aagttaagc tgtgttaaatt    780
gatgaacgat atcagtcctt caaactcttt acttctgttt ttggagtggg actgaagaca    840
tctgagaaat ggttcaggat ggggttcaga tctctgagta aaataatgtc agacaaaacc    900
ctgaaattca caaaaatgca gaaagcagga tttctctatt atgaagacct tgtcagctgc    960
gtgaccaggg ccgaagcaga ggcggttggc gtgctggtta aagaggctgt gtgggcattt   1020
ctgccggatg cctttgtcac catgacagga ggattccgca ggggtaagaa gattgggcat   1080

```

-continued

```

gatgtagatt ttttaattac cagcccagga tcagcagagg atgaagagca acttttgcct 1140
aaagtgataa acttatggga aaaaaaggga ttacttttat attatgacct tgtggagtca 1200
acatttgaaa agttcaagtt gccaagcagg caggtggata ctttagatca ttttcaaaaa 1260
tgctttctga ttttaaaatt gcaccatcag agagtagaca gtagcaagtc caaccagcag 1320
gaaggaaaga cctggaaggc catccgtgtg gacctgggta tgtgccccta cgagaaccgt 1380
gcctttgccc tgctaggctg gactggctcc cggcagtttg agagagacat ccggcgctat 1440
gccacacacg agcgggaagat gatgctggat aaccacgctt tatatgacaa gaccaagagg 1500
gtattttctca aagcggaag tgaagaagaa atctttgcac atctgggatt ggactacatt 1560
gaaccatggg aaagaaatgc ttaggagaaa gctgtcaact tttttctttt ctgttctttt 1620
tttcagggtta gacaaattat gcttcatatt ataataagaa atgccttagt caagtttggg 1680
attctttaca ttttaccag atgtagattg cttctagaaa taagtagttt tggaaacgtg 1740
atcaggcacc ccctgggtta tgctctggca agccatttgc aggactgatg tgtagaactc 1800
gcaatgcatt ttccatagaa acagtgttgg aattgggtggc tcatttccag ggaagttcat 1860
caaagccac tttgccaca gtgtagctga aatactgtat acttgccaat aaaaatagga 1920
aac 1923

```

```

<210> SEQ ID NO 3
<211> LENGTH: 465
<212> TYPE: PRT
<213> ORGANISM: Escherichia coli

```

```

<400> SEQUENCE: 3

```

```

Met Phe Thr Arg Val Ala Asn Phe Cys Arg Lys Val Leu Ser Arg Glu
1           5           10          15
Glu Ser Glu Ala Glu Gln Ala Val Ala Arg Pro Gln Val Thr Val Ile
20          25          30
Pro Arg Glu Gln His Ala Ile Ser Arg Lys Asp Ile Ser Glu Asn Ala
35          40          45
Leu Lys Val Met Tyr Arg Leu Asn Lys Ala Gly Tyr Glu Ala Trp Leu
50          55          60
Val Gly Gly Gly Val Arg Asp Leu Leu Leu Gly Lys Lys Pro Lys Asp
65          70          75          80
Phe Asp Val Thr Thr Asn Ala Thr Pro Glu Gln Val Arg Lys Leu Phe
85          90          95
Arg Asn Cys Arg Leu Val Gly Arg Arg Phe Arg Leu Ala His Val Met
100         105         110
Phe Gly Pro Glu Ile Ile Glu Val Ala Thr Phe Arg Gly His His Glu
115         120         125
Gly Asn Val Ser Asp Arg Thr Thr Ser Gln Arg Gly Gln Asn Gly Met
130         135         140
Leu Leu Arg Asp Asn Ile Phe Gly Ser Ile Glu Glu Asp Ala Gln Arg
145         150         155         160
Arg Asp Phe Thr Ile Asn Ser Leu Tyr Tyr Ser Val Ala Asp Phe Thr
165         170         175
Val Arg Asp Tyr Val Gly Gly Met Lys Asp Leu Lys Asp Gly Val Ile
180         185         190
Arg Leu Ile Gly Asn Pro Glu Thr Arg Tyr Arg Glu Asp Pro Val Arg

```

-continued

195	200	205
Met Leu Arg Ala Val Arg Phe Ala Ala Lys Leu Gly Met Arg Ile Ser		
210	215	220
Pro Glu Thr Ala Glu Pro Ile Pro Arg Leu Ala Thr Leu Leu Asn Asp		
225	230	235 240
Ile Pro Pro Ala Arg Leu Phe Glu Glu Ser Leu Lys Leu Leu Gln Ala		
	245	250 255
Gly Tyr Gly Tyr Glu Thr Tyr Lys Leu Leu Cys Glu Tyr His Leu Phe		
	260	265 270
Gln Pro Leu Phe Pro Thr Ile Thr Arg Tyr Phe Thr Glu Asn Gly Asp		
	275	280 285
Ser Pro Met Glu Arg Ile Ile Glu Gln Val Leu Lys Asn Thr Asp Thr		
	290	295 300
Arg Ile His Asn Asp Met Arg Val Asn Pro Ala Phe Leu Phe Ala Ala		
305	310	315 320
Met Phe Trp Tyr Pro Leu Leu Glu Thr Ala Gln Lys Ile Ala Gln Glu		
	325	330 335
Ser Gly Leu Thr Tyr His Asp Ala Phe Ala Leu Ala Met Asn Asp Val		
	340	345 350
Leu Asp Glu Ala Cys Arg Ser Leu Ala Ile Pro Lys Arg Leu Thr Thr		
	355	360 365
Leu Thr Arg Asp Ile Trp Gln Leu Gln Leu Arg Met Ser Arg Arg Gln		
	370	375 380
Gly Lys Arg Ala Trp Lys Leu Leu Glu His Pro Lys Phe Arg Ala Ala		
385	390	395 400
Tyr Asp Leu Leu Ala Leu Arg Ala Glu Val Glu Arg Asn Ala Glu Leu		
	405	410 415
Gln Arg Leu Val Lys Trp Trp Gly Glu Phe Gln Val Ser Ala Pro Pro		
	420	425 430
Asp Gln Lys Gly Met Leu Asn Glu Leu Asp Glu Glu Pro Ser Pro Arg		
	435	440 445
Arg Arg Thr Arg Arg Pro Arg Lys Arg Ala Pro Arg Arg Glu Gly Thr		
	450	455 460
Ala		
465		

<210> SEQ ID NO 4

<211> LENGTH: 1398

<212> TYPE: DNA

<213> ORGANISM: Escherichia coli

<400> SEQUENCE: 4

```

attttttacc gagtcgctaa tttttgccgc aaggtgctaa gccgcgagga aagcgaggct    60
gaacaggcag tcgcccgtcc acaggtgacg gtgatcccgc gtgagcagca tgctatttcc    120
cgcaaagata tcagtaaaa tgccctgaag gtaatgtaca ggctcaataa agcgggatac    180
gaagcctggc tggttggcgg cggcgtgcgc gacctgttac ttggcaaaaa gccgaaagat    240
tttgacgtaa ccactaacgc cacgcctgag caggtgcgca aactgttcgg taactgccgc    300
ctggtggggtc gccgtttccg tctggctcat gtaatgtttg gcccgagat tatcgaagtt    360
gcgaccttcc gtggacacca cgaaggtaac gtcagcgacc gcacgacctc ccaacgcggg    420
caaaacggca tgttgctgcg cgacaacatt ttcggctcca tcgaagaaga cgcccagcgc    480

```

-continued

```

cgcgatttca ctatcaacag cctgtattac agcgtacggg attttaccgt ccgtgattac 540
gttgggcgga tgaaggatct gaaggacggc gttatccgtc tgattggtaa cccggaacg 600
cgctaccgtg aagatccggt acgtatgctg cgcgcgggtac gttttgccgc caaattgggt 660
atgcgcatca gcccggaaac cgcagaaccg atccctcgcc tcgtaccct gctgaacgat 720
atcccaccgg cagcctgtgt tgaagaatcg cttaaaactgc tacaagcggg ctacggttac 780
gaaacctata agctgttgtg tgaatatcat ctgttcacgc cgctgttccc gaccattacc 840
cgctacttca cggaataatgg cgacagcccg atggagcgga tcattgaaca ggtgctgaag 900
aataccgata cgcgtatcca taacgatatg cgcgtgaacc cggcgttccg gtttgccgcc 960
atgttctggt acccactgct ggagacggca cagaagatcg cccaggaaag cggcctgacc 1020
tatcacgacg ctttcgcgct ggcgatgaac gacgtgctgg acgaagcctg ccgttcactg 1080
gcaatcccg aacgtctgac gacattaacc cgcgatatct ggagttgca gttgcgtatg 1140
tcccgctgct agggtaaacc cgcattgaaa ctgctggagc atcctaagtt ccgtgcggct 1200
tatgacctgt tggccttgcg agctgaagtt gagcgtaacg ctgaactgca gcgtctggtg 1260
aaatggtggg gtgagttcca ggttccgcg ccaccagacc aaaaagggat gctcaacgag 1320
ctggatgaag aaccgtcacc gcgtcgtcgt actcgtcgtc cagcgaacg cgcaccacgt 1380
cgtgagggta ccgcatga 1398

```

```

<210> SEQ ID NO 5
<211> LENGTH: 349
<212> TYPE: PRT
<213> ORGANISM: Escherichia coli

```

```

<400> SEQUENCE: 5

```

```

Gly Ser His Met Ser Tyr Gln Lys Val Pro Asn Ser His Lys Glu Phe
1      5      10      15
Thr Lys Phe Cys Tyr Glu Val Tyr Asn Glu Ile Lys Ile Ser Asp Lys
20     25     30
Glu Phe Lys Glu Lys Arg Ala Ala Leu Asp Thr Leu Arg Leu Cys Leu
35     40     45
Lys Arg Ile Ser Pro Asp Ala Glu Leu Val Ala Phe Gly Ser Leu Glu
50     55     60
Ser Gly Leu Ala Leu Lys Asn Ser Asp Met Asp Leu Cys Val Leu Met
65     70     75     80
Asp Ser Arg Val Gln Ser Asp Thr Ile Ala Leu Gln Phe Tyr Glu Glu
85     90     95
Leu Ile Ala Glu Gly Phe Glu Gly Lys Phe Leu Gln Arg Ala Arg Ile
100    105    110
Pro Ile Ile Lys Leu Thr Ser Asp Thr Lys Asn Gly Phe Gly Ala Ser
115    120    125
Phe Gln Cys Asp Ile Gly Phe Asn Asn Arg Leu Ala Ile His Asn Thr
130    135    140
Leu Leu Leu Ser Ser Tyr Thr Lys Leu Asp Ala Arg Leu Lys Pro Met
145    150    155    160
Val Leu Leu Val Lys His Trp Ala Lys Arg Lys Gln Ile Asn Ser Pro
165    170    175
Tyr Phe Gly Thr Leu Ser Ser Tyr Gly Tyr Val Leu Met Val Leu Tyr
180    185    190

```

-continued

Tyr Leu Ile His Val Ile Lys Pro Pro Val Phe Pro Asn Leu Leu Leu
 195 200 205
 Ser Pro Leu Lys Gln Glu Lys Ile Val Asp Gly Phe Asp Val Gly Phe
 210 215 220
 Asp Asp Lys Leu Glu Asp Ile Pro Pro Ser Gln Asn Tyr Ser Ser Leu
 225 230 235 240
 Gly Ser Leu Leu His Gly Phe Phe Arg Phe Tyr Ala Tyr Lys Phe Glu
 245 250 255
 Pro Arg Glu Lys Val Val Thr Phe Arg Arg Pro Asp Gly Tyr Leu Thr
 260 265 270
 Lys Gln Glu Lys Gly Trp Thr Ser Ala Thr Glu His Thr Gly Ser Ala
 275 280 285
 Asp Gln Ile Ile Lys Asp Arg Tyr Ile Leu Ala Ile Glu Asp Pro Phe
 290 295 300
 Glu Ile Ser His Asn Val Gly Arg Thr Val Ser Ser Ser Gly Leu Tyr
 305 310 315 320
 Arg Ile Arg Gly Glu Phe Met Ala Ala Ser Arg Leu Leu Asn Ser Arg
 325 330 335
 Ser Tyr Pro Ile Pro Tyr Asp Ser Leu Phe Glu Glu Ala
 340 345

<210> SEQ ID NO 6
 <211> LENGTH: 1047
 <212> TYPE: DNA
 <213> ORGANISM: Escherichia coli

<400> SEQUENCE: 6

```

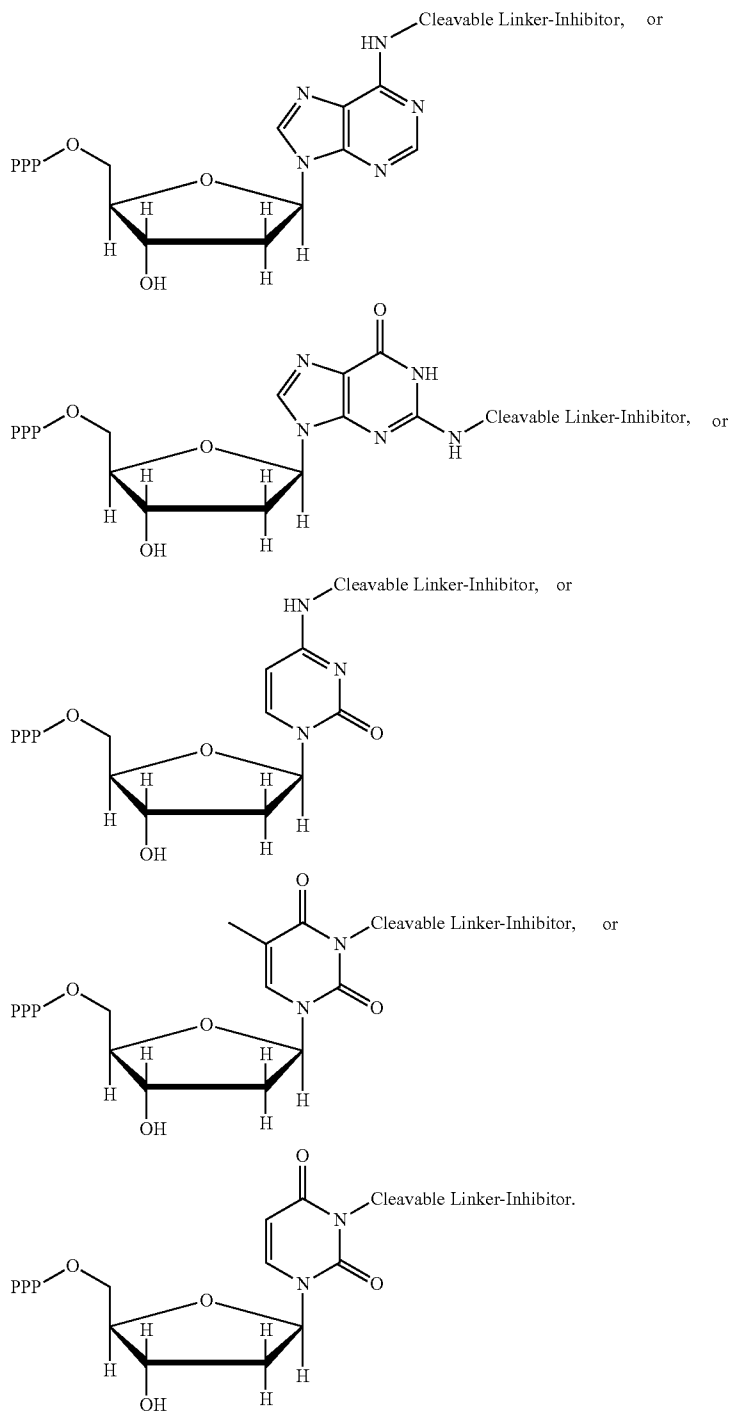
ggcagccata tgagctatca gaaagtgccg aacagccata aagaatttac caaattttgc      60
tatgaagtgt ataacgaaat taaaattagc gataaagaat ttaaagaaaa acgcgcggcg      120
ctggataccc tgcgcctgtg cctgaaacgc attagcccg gtcggaact ggtggcggtt      180
ggcagcctgg aaagcggcct ggcgctgaaa aacagcgata tggatctgtg cgtgctgatg      240
gatagccgcg tgcagagcga taccattgcg ctgcagtttt atgaagaact gattgcggaa      300
ggctttgaag gcaaatttct gcagcgcgcg cgcattccga ttattaaact gaccagcgat      360
accaaaaacg gctttggcgc gagctttcag tgcgatattg gctttaacaa ccgcctggcg      420
attcataaca ccctgctgct gagcagctat accaaactgg atgcgcgcct gaaaccgatg      480
gtgctgctgg tgaacattg ggcgaaacgc aaacagatta acagcccgta ttttggcacc      540
ctgagcagct atggctatgt gctgatgggt ctgtattatc tgattcatgt gattaaaccg      600
ccggtgtttc cgaacctgct gctgagcccg ctgaaacagg aaaaaattgt ggatggcgtt      660
gatgtgggct ttgatgataa actggaagat attccgccga gccagaacta tagcagcctg      720
ggcagcctgc tgcattggct ttttcgcttt tatgcgtata aatttgaacc gcgcgaaaaa      780
gtggtgacct ttgcgcgcc ggatggctat ctgaccaaac aggaaaaagg ctggaccagc      840
gcgaccgaac ataccggcag cgcggatcag attattaaag atcgctatat tctggcgatt      900
gaagatccgt ttgaaattag ccataacgtg ggccgcaccg tgagcagcag cggcctgtat      960
cgcatcgcg gcgaatttat ggccgcgagc cgctgctga acagccgcag ctatccgatt      1020
ccgtatgata gcctgtttga agaagcg                                     1047
  
```

1. A method for synthesizing an oligonucleotide, comprising:

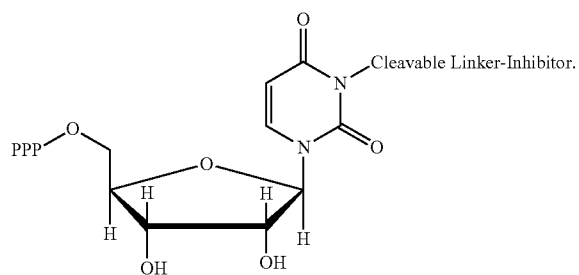
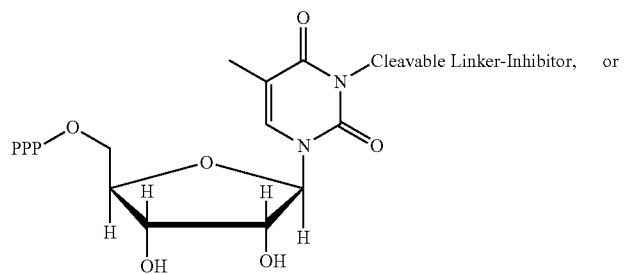
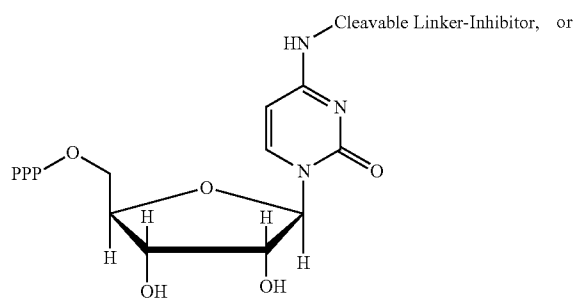
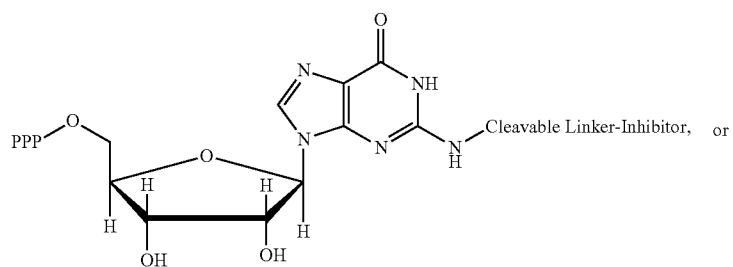
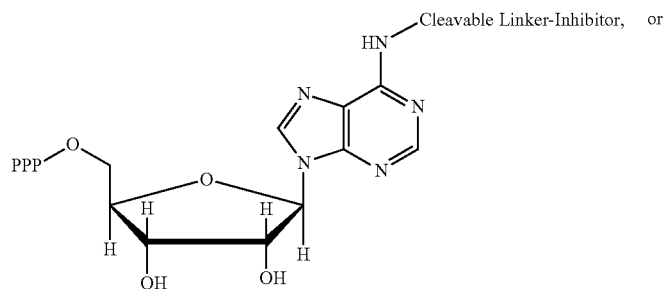
exposing an oligonucleotide attached to a solid support to a nucleotide analog in the presence of a nucleotidyl transferase enzyme and in the absence of a nucleic acid template such that the nucleotide analog is incorporated into the oligonucleotide,

wherein the nucleotide analog comprises a nucleotide coupled, by a cleavable linker, to an inhibitor that prevents the nucleotidyl transferase from catalyzing incorporation of a nucleotide or an additional nucleotide analog into said oligonucleotide until said inhibitor is removed by cleavage of said cleavable linker.

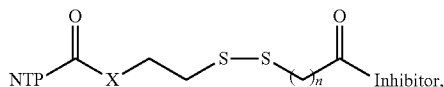
2. The method of claim 1, wherein the nucleotide analog comprises the following structure:



3. The method of claim 1, wherein the nucleotide analog comprises the following structure:

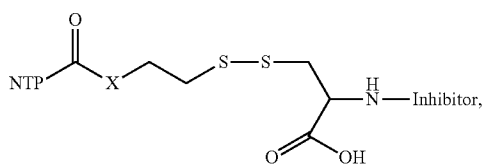


4. The method of claim 1, wherein the nucleotide analog comprises the following cleavable linker structure:



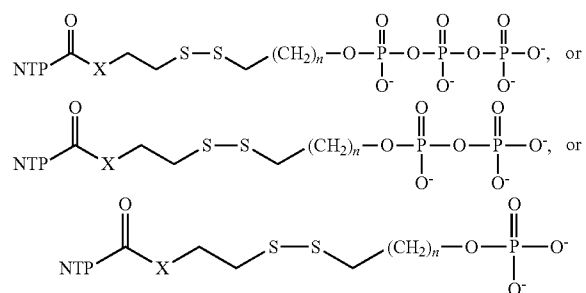
wherein NTP=nucleotide triphosphate, $n=1, 2, 3$, or 4 , and $-\text{X}-$ is $-\text{O}-$, $-\text{S}-$, $-\text{NH}-$, or $-\text{CH}_2-$.

5. The method of claim 1, wherein the nucleotide analog has the following cleavable linker structure:



wherein NTP=nucleotide triphosphate, $-\text{X}-$ is $-\text{O}-$, $-\text{S}-$, $-\text{NH}-$, or $-\text{CH}_2-$.

6. The method of claim 1, wherein the nucleotide analog has the following inhibitor structure:



wherein NTP=nucleotide triphosphate, $n=1, 2, 3$, or 4 , and $-\text{X}-$ is $-\text{O}-$, $-\text{S}-$, $-\text{NH}-$, or $-\text{CH}_2-$.

7. The method of claim 1, wherein the nucleic acid attached to the solid support is exposed to the nucleotide analog in the presence of an aqueous solution having a pH between about 6.5 and 8.5.

8. The method of claim 1, the nucleic acid attached to the solid support is exposed to the nucleotide analog in the presence of an aqueous solution at a temperature between about 35 and 39° C.

9. The method of claim 1, wherein the solid support is a bead, a well, or a peg.

10. The method of claim 1, wherein the nucleic acid is single stranded.

11. The method of claim 1, wherein the cleavable linker comprises a moiety that forms a cyclic by-product when cleaved from the nucleotide analog.

12. The method of claim 1, further comprising:

cleaving the cleavable linker in order to produce a native nucleotide; and

exposing the native nucleotide to a second nucleotide analog in the presence of a nucleotidyl transferase enzyme and in the absence of a nucleic acid template.

13. The method of claim 1, further comprising providing an aqueous solution comprising the nucleotide analog and the nucleotidyl transferase enzyme.

14. The method of claim 1, wherein the inhibitor comprises a charged moiety.

15. The method of claim 14, wherein the charged moiety comprises a negative charge.

16. The method of claim 15, wherein the inhibitor comprises an amino acid with a negative charge.

17. The method of claim 16, wherein the amino acid is acetylated.

18. The method of claim 14, wherein the charged moiety comprises a positive charge.

19. The method of claim 18, wherein the inhibitor comprises an amino acid with a positive charge.

20. The method of claim 19, wherein the amino acid is acetylated.

21. The method of claim 1, wherein the nucleotide analog comprises a ribose sugar or a deoxyribose sugar.

22. The method of claim 1, wherein the nucleotide substrate comprises a base selected from the group consisting of adenine, guanine, cytosine, thymine, and uracil.

23. The method of claim 1, wherein the nucleotidyl transferase comprises a protein sequence that is at least about 90% identical to SEQ ID NO. 1, SEQ ID NO. 3, or SEQ ID NO. 5.

24. The method of claim 1, wherein the nucleotidyl transferase originates from an organism having a nucleotide sequence that is at least about 90% identical to SEQ ID NO. 2, SEQ ID NO. 4, or SEQ ID NO. 6.

25. The method of claim 1, wherein the nucleotide analog is encapsulated in a droplet.

26. The method of claim 1, wherein the nucleotidyl transferase enzyme is encapsulated in a droplet.

* * * * *